

Universidade do Minho
Escola de Engenharia

Morgana Carneiro de Andrade

Interoperabilidade semântica entre repositórios digitais científicos: um perfil de aplicação de metadados baseado nas necessidades de informação em oncologia

Morgana Carneiro de Andrade
Interoperabilidade semântica entre repositórios digitais científicos: um perfil de aplicação de metadados baseado nas necessidades de informação em oncologia

UMinho | 2018

dezembro de 2018

Universidade do Minho
Escola de Engenharia

Morgana Carneiro de Andrade

Interoperabilidade semântica entre repositórios digitais científicos: um perfil de aplicação de metadados baseado nas necessidades de informação em oncologia

Tese de Doutoramento
Programa Doutoral Tecnologias e Sistemas de Informação

Trabalho efetuado sob a orientação da
Professora Doutora Ana Alice Baptista

dezembro de 2018

DECLARAÇÃO DE INTEGRIDADE

Declaro ter atuado com integridade na elaboração da presente tese. Confirmo que em todo o trabalho conducente à sua elaboração não recorri à prática de plágio ou a qualquer forma de falsificação de resultados.

Mais declaro que tomei conhecimento integral do Código de Conduta Ética da Universidade do Minho.

Universidade do Minho,_____

Nome completo: Morgana Carneiro de Andrade

Assinatura: 

Agradeço

> [...] todas as dificuldades que enfrentei; não fosse por elas, eu não teria saído do lugar. As facilidades nos impedem de caminhar [...].
>
> Chico Xavier

À Professora Ana Alice Baptista, pela oportunidade, pelo aprendizado, pela amizade e, sobretudo, pela confiança. E o que dizer da simplicidade com que expõe seus conhecimentos e a generosidade em compartilhá-lo? Qualidades que são exemplos de conduta na vida.

A todo corpo Docente da Universidade do Minho, nas figuras do Prof. João Álvaro Carvalho, Profa. Izabel Ramos, Profa. Delfina Soares e Prof. João Carlos Nascimento.

Aos membros do jurí pelas valiosas contribuições.

Aos "meninos" do laboratório, Jorge e Rui; à Maria João e à Manuela, da Secretaria do PDTSI.

Ao meu grande amigo e colega Victor Barros. Obrigada por tudo, por ouvir e por dividir grandes momentos!

A dois amigos que foram tão presentes nesse caminhar: Fernanda Stefens, Nelson Freitas e Fabio Longo de Moura.

Às minhas amigas e aos amigos /colegas de doutoramento, Isabel Ferreira, Cândida, Antonio e José Luis (a ponte para meus investigados). Foram momentos únicos e inesquecíveis.

À Mariana Malta, obrigada pela presença (à distância) em todo o trabalho, mesmo sem perceber o quanto me apoiou. Ainda me recordo do seu apoio ao sair da minha qualificação!

Àqueles que participaram desta investigação, direta ou indiretamente. Sem vocês, eu não teria chegado aqui. Meu muito obrigada à Dra. Carla Oliveira, IPATIMUP; Rita de Cássia Borges, Sírio Libanês; Prof. Edmundo Luis Pereira, UFPA, Nelma Fezer e demais participantes.

À Universidade do Minho, onde realizei o doutoramento; à Universidade Federal do Espírito Santo, onde exerço minha profissão; à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – CAPES, pela oportunidade de voar mais alto, de buscar novos horizontes e conhecimento.

Ao professor José Geraldo Mill, por sempre ter acreditado em mim, meu muito obrigada!

À Raquel Madeira, Arlete Franco e Maria Lúcia Cassatti, pela confiança depositada.

À Alina, que, ao longo da estrada que me trouxe até aqui, contribuiu de forma ímpar para minhas vírgulas e meus pontos.

Impossível incluir todos aqueles que contribuíram para que eu alcançasse meus ideais. Foi uma estrada muito longa, com muitas paradas, umas à sombra e outras sob sol ou chuva. Mesmo que eu não o(a) cite, caso você tenha acesso a esse trabalho, saberá o quanto eu sou grata e o porquê. Por outro lado, é irresistível não registrar:

Aos Freitas, por ter me acolhido, como se da família fosse!

Aos meus amigos em Portugal, que trago na lembrança e no coração, Nelson Freitas, Cecília e Manel, Naike e Nuno, Fernanda e Zezinho, e Francklin. Muito obrigada por fazer me sentir parte deste país. À Maria José Baños (Jojo), pela sincronia e alegria. Ceren and family, "for engagement and kindness".

Ainda em Portugal, Andrômeda Menezes, Rayanne Souza, Fran e Júlia Brasil, pela presença e pelos momentos alegres!

No Brasil, às minhas amigas Maria Luíza Fontenelle Dumans, um exemplo como educadora, profissional e amiga, e à Profa. Brígida Cervantes, minha orientadora no Mestrado, maior e primeira incentivadora para que eu ingressasse no doutoramento.

À Angela Maria Becalli, por todas as oportunidades que criou na minha vida, pela amizade e pela confiança!

À Gleice Pereira, colega e parceira! Por mais que as nuvens encubram o sol, sabemos que ele sempre estará lá. Obrigada!

À Patrícia Barros, pela paciência, o saber ouvir e a parceria.

A Gutemberg Brasil, pelo diálogo com os números.

À Taciana Lemos, pelo apoio.

A Fábio Medina, diretor da Biblioteca Central da Ufes, pela compreensão!

Ao Fabio e a Nathaniely por todo o apoio ao fim do percurso.

À Minha família, cuja "semântica" é inquestionável! Vocês são tudo para mim! Obrigada minha Mãe por ser tão teimosa, tão guerreira e tão presente! Quanto desejei ter o prazer de dizer para a senhora e para meu pai, esteja onde estiver, que eu consegui! Muito obrigada pela vida que vocês me proporcionaram. Obrigada meus irmãos e minha irmã por nós sermos uma FAMÍLIA!

Dedico

A

minha **M**ãe, por ser quem é, pelo amor que tem à vida, por sua luta e seu exemplo.
minha **i**rmã, por sua força, seu amor, seu apoio e por acreditar sempre!
Jú**n**ior, Max e Elmanno, cada um com seu jeito de amar e de ser!
hoje e sempre,
P**a**i, pelas lembranças e por todo o amor!

A todos que **f**azem parte da minha história!
A ti**a**s, primas e primos, que tanto amo,
Aos **m**eus sobrinhos e sobrinhas.
A Luc**i**ano, por ser tão "filho", "irmão" e companheiro.
À C**l**eide, pelo carinho.
Aos Am**i**gos, aos de cá e aos do além-mar!
E **a** todos que lutam pela Vida!

É meu e vosso este fado
Destino que nos amarra
Por mais que seja negado
Às cordas de uma guitarra

Sempre que se ouve o gemido
De uma guitarra a cantar
Fica-se logo perdido
Com vontade de chorar

Ó gente da minha terra
Agora é que eu percebi
Esta tristeza que trago
Foi de vós que recebi

E pareceria ternura
Se eu me deixasse embalar
Era maior a amargura
Menos triste o meu cantar

Ó gente da minha terra
Agora é que eu percebi
Esta tristeza que trago
Foi de vós que recebi.

(Amália Rodrigues)

Por isso na impaciência
Desta sede de saber,
Como as aves do deserto
As almas buscam beber...
Oh! Bendito o que semeia
Livros... livros à mão cheia...
E manda o povo pensar!
O livro caindo n'alma
É germe – que faz a palma,
É chuva – que faz o mar.

Castro Alves

RESUMO

Andrade M. C. (2018). *Interoperabilidade semântica entre repositórios digitais científicos: um perfil de aplicação de metadados baseado nas necessidades de informação em Oncologia*. Tese de Doutorado, Programa de Doutoramento em Tecnologias e Sistemas de Informação, Universidade do Minho, Guimarães, Portugal.

Introdução: O acesso à informação no domínio da Oncologia tem demonstrado ser fundamental para as investigações e o atendimento médico, notadamente as informações científicas armazenadas em sistemas de informação, a exemplo das bases de dados bibliográficas e dos repositórios digitais científicos. Os repositórios, embora consolidados como alternativa para o armazenamento e o acesso livre às publicações científicas, ainda não atingiram um nível de interoperabilidade semântica que assegure uma descoberta eficaz dos recursos ali armazenados. A interoperabilidade semântica entre os sistemas de informação requer o estabelecimento de padrões, especialmente em relação à descrição dos recursos e seus conteúdos. Nesse sentido, há recomendações que esses padrões, assim como regras e elementos de metadados, sejam descritos por meio de perfis de aplicação de metadados e, posteriormente, disponibilizados por esses sistemas. Este estudo se refere às boas práticas dos metadados, que inclui o perfil de aplicação, à interoperabilidade semântica, às diretrizes *Linked Data*, aos repositórios digitais científicos e às necessidades de informação de investigadores e médicos no domínio da Oncologia. **Objetivo**: Identificar tipos de necessidades de informação de médicos e investigadores no domínio da Oncologia, de modo a subsidiar o desenvolvimento de um perfil de aplicação de metadados para repositórios digitais científicos. **Metodologia**: Foi adotado o Design Science Research, que incluiu o desenvolvimento de análises sistemáticas de documentos e publicações científicas, obtidas por meio das pesquisas bibliográfica e documental (documentos obtidos em *sites* e bases de dados bibliográficas); *survey* com administradores de repositórios digitais científicos, para identificação de perfis de aplicação, esquemas e elementos de metadados utilizados por repositórios digitais científicos; estudo de caso (aplicação de *survey* entre médicos e investigadores em Portugal, para a identificação de tipos de necessidades de informação); e a técnica Delphi, realizada em duas rondas, para validação dos resultados que integram o perfil de aplicação de metadados concebido a partir de tipos de necessidades de informação. Esses tipos de necessidades de informação foram traduzidos em classes e propriedades submetidas à análise de bibliotecários que atuam em recuperação, tratamento, organização e disseminação da informação no domínio da Oncologia. **Conclusão**: Os estudos realizados no contexto desta tese convergiram para a

identificação de tipos de necessidades informacionais associadas aos médicos e investigadores que atuam no domínio da Oncologia, que proporcionou um conjunto de elementos de metadados utilizados no desenvolvimento de um perfil de aplicação de metadados voltado para repositórios digitais científicos (Onco-MAP). O perfil de aplicação, além de incluir elementos de esquemas de metadados já existentes, exigiu a criação de novo esquema de metadados, denominado Onco.

Palavras-chave: Pefil de aplicação de metadados. Oncologia. Repositório Digital Científico. Tipos de necessidades de informação. *Linked Open Data*.

ABSTRACT

Andrade M. C. (2018). *Semantic interoperability between scientific digital repositories: a metadata application profile based on information needs in Oncology*. Doctoral Thesis, Information Systems and Technologies Doctorate Program, Minho University, Guimarães, Portugal.

Introduction: access to information in Oncology domain has proven fundamental for researches and treatments, notably the scientific information stored in information systems, such as databases and scientific digital repositories. The repositories, although consolidated as alternative to storage and free access to scientific publications, have not yet reached a semantic interoperability level that assures an effective discovery of resources stored there. The semantic interoperability among the information systems requires the establishment of standards, particularly regarding the description of the resources and their contents. In this respect, there are recommendations that these standards, as well as rules and metadata elements, are described through metadata application profiles and made available for these systems. This study refers to the good practices of the metadata, which includes the application profile, semantic interoperability, the Linked Data guidelines, the scientific digital repositories and the information needs of researchers and physicians in the Oncology domain. **Aim**: to identify types of information needs of physicians and researchers in the Oncology domain, in order to subsidize the development of a metadata application profile for scientific digital repositories. **Methodology**: It was employed the Design Science Research, which included the development of systematic analysis of documents and scientific publications obtained through bibliographic and documentary searches (documents obtained on sites and bibliographic databases); survey with managers of scientific digital repositories for identification of application profiles, schemes and metadata elements used for scientific digital repositories; case study (application of survey between physicians and researchers in Portugal for identifying types of information needs); and the Delphi technique performed in two rounds for validation of results within the metadata application profile designed from types of information needs. These types of information needs were translated into classes and properties submitted to analysis of librarians who work in recovery, treatment, organization and dissemination of information in the Oncology domain. **Conclusion**: The studies carried out in the context of this thesis converged for the identification of types of information needs associated with the physicians and researchers who work in the Oncology domain, which provided a set of metadata elements used in development of a metadata application profile addressed to scientific digital repositories (Onco-MAP). The application profile,

in addition to include elements of existing metadata schemes, required creation of a new metadata scheme, named Onco.

Keywords: Metadata application profile. Oncology. Scienfic Digital Repository. Types of information needs. *Linked Open Data*.

LISTA DE ABREVIAÇÕES E SIGLAS

AACR2 – Anglo-American Cataloguing Rules
AIDS/SIDA – Síndrome da Imunodeficiência Adquirida
AMF – Academic Metadata Format
BBC – British Broadingcasting Corporation
BDB – Base de Dados Bibliográficas
BIOSIS – BioSciences Information Services
CAPES – Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior
CAS – Chemical Abstracts Services
CBAC – Chemical and Biological Activities
CDISC – Quality Controlled Programming & Design Reporting Analysis of Research
CNPq – Conselho Nacional de Desenvolvimento Científico e Tecnológico
DARE – Digital Academic Repository
DC – Dublin Core
DCAM – Dublin Core Abstract Model
DCAP – Dublin Core Application Profile
DCLib – DC Library Application Profile
DCMES – Dublin Core Metadata Element Set
DCMI – Dublin Core Metadata Initiative
DCMI Terms – DCMI Metadata Terms
DIM – DSPACE intermediate metadata
DNA – Ácido Desoxirribonucleico
DOAJ – Directory of Open Access Journals
DOI – Digital Object Identifier
DRIVER – Digital Repository Infrastructure Vision for European Research
DSP – Dublin Core Description Set Profile
DSR – Design Science Research
EAD – Encoded Archival Description
EBM – Medicina Baseada em Evidências
EDM – Europeana Data Model
EIF – European Interoperability Framework
EMBASE – Excerpta Medica Database
ERIC – Education Resources information Center

ETD-MS – Electronic Theses and Dissertations Metadata Standard

ETD-MS – Eletronic Theses and Dissertations Metadata Standard

EUA – Estados Unidos da América

EUA – Estados Unidos da América

FAO – Food and Agriculture Organization

FOAF – Friend of a Friend

FDA – United States Food and Drug Administration

FRBR – Requisitos Funcionais para Registros Bibliográficos

HL7 – Health Level Seven International

HTML – Hyper Text Markup Language

HTTP – Hypertext Transfer Protocol

I&D – Investigação e Desenvolvimento

IBMC – Instituto de Biologia Molecular e Celular

ICD – Internation Classification of Desease

ID – Identificadores Persistentes

IDABC – European e-Government Services

IEEE – Institute of Electrical and Electronics Engineers

INCA – Instituto Nacional do Câncer

INEB – Instituto de Engenharia Biomédica

IPATIMUP – Instituto de Patologia e Imunologia Molecular da Universidade do Porto

IPO – Instituto Português de Oncologia

IRI – Identificador de Recursos Internacionalizados

ISO – International Organization for Standardization

ISSN – International Standard Serial Number

JISC – Joint Information Science Committee

KOS – Sistemas de Organização do Conhecimento

LCSH – Library Congress Subject Heading

LISA – Library and Information Science Abstracts

LISTA – Library, Information Science and Technology Abstract

LOD – *Linked Open Data*

LOM – Learning Object Metadata

LOV – Language open Vocabulary

MARC – Machine-Readable Cataloguing

MEDLARS – Medical Literature Analysis and Retrieval System

MEDLINE – Medical Literature Analysis and Retrieval System Online
MESH – Medical Subject Heading
METS – Metadata Encoding and Transmission Standard
MODS – Metadata Object Description Schema
MPEG – Moving Picture Experts Group
NCI – Instituto Nacional de Câncer of the United States
NI – Necessidade de informação
NISO – National Information Standards Organizations
NLM – National Library of Medicine
NSDL – National Science Digital Library
OAI – Open Archive Iniciative
OAI-DC – Open Archives Initiative Dublin Core
OAI-ORE – Open Archives Initiative - Object Reuse and Exchange
OAI-PMH – Archive Iniciative Protocol for Metada Haversting
OWL – Web Ontology Language
PA – Perfil de Aplicação
PAM – Perfil de Aplicação de Metadados
PICO – Problema, Intervenção, Comparação e Resultado
PREMIS – Preservation Metadata: implementation Strategies
PRISM – Publishing Requirements for Industry Standard Metadata
RDC – Repositórios Digitais Científicos
RDF – Resource Description Framework
RDFS – Resource Description Framework Schema
RI – Repositório Institucional
RO – Repositório Organizacional
ROE – Repositório de Objetos de Ensino
RT – Repositório Temático
RTD – Repositório de Teses e Dissertações
SAP – Subject Access Point
SES – Syntax Encoding Scheme
SHACL – Shapes Constraint Language
SKOS – Simple Knowledge Organization Schemes
SOC – Sistema de Organização do Conhecimento
SNOMED-CT – SNOMED Clinical Terms

SPARQL – SPARQL Protocol and RDF Query Language
SPSS – Statistical Package for the Social Sciences
SWAP – Scholarly Work Application Profile
TBITS – Treasury Board Information Management Standard
TI – Tecnologia da Informação
TIC – Tecnologia da Informação e Comunicação
TNI – Tipo de Necessidades de Informação
UE – União Europeia
UKOLN – United Kingdom Office for Library and Information Networking
UML – Unified Modelling Language
UMLS – Unified Medical Language System
URI – Identificador Único de Recurso
URL – Uniform Resource Locator
VES – Vocabulary Encoding Scheme
VIAF – Virtual International Authority File
VOA3R – The Virtual Open Access Agriculture & Aquaculture Repository
W3C – World Wide Web Consortium
WOS – Web of Science
WWW – World Wide Web
XML – Extensible Markup Language (Linguagem de Descrição Específica)
XMLS – Extensible Markup Language Schema

LISTA DE FIGURAS

Figura 1. Estrutura da tese. 11
Figura 2. *Layer Cake* da Web Semântica 15
Figura 3. Triplos RDF para descrição do recurso Aspirina 16
Figura 4. Níveis de interoperabilidade Dublin Core. 24
Figura 5. Fazer investigação, comunicar e aplicar os resultados - diagrama do contexto.. 37
Figura 6. Fazer a pesquisa, comunicação e aplicação de resultados: desagregação (A0). 38
Figura 7. Performance da pesquisa (A2) 39
Figura 8. Facilitar a recuperação global (A3221). 42
Figura 9. Fazer manuscrito ou cópia da publicação em acesso aberto (A32212). 44
Figura 10. Integração de metadados nos serviços (A32213).. 45
Figura 11. Conheça a publicação (A3231) 46
Figura 12. Leitura da publicação (A32342). 49
Figura 13. Aplicação na prática (A4) 49
Figura 14. Ciclos do Design Science Research 62
Figura 15. Design Science Research adotado para o desenvolvimento do perfil de aplicação.. 63
Figura 16. Identificação de Perfis de Aplicação de Metadados. 69
Figura 17. Tipos de materiais armazenados nos RDCs. 124
Figura 18. Esquema de metadados utilizados por RDCs. 126
Figura 19. Relacionamento entre a clínica e a investigação básica.. 161
Figura 20. Representação gráfica do processo de pesquisa bibliográfica. 167
Figura 21. Nuvem de palavras com os termos mais frequentes nos artigos selecionados. 170
Figura 22. Nuvem de palavras com TNIs dos médicos. 172
Figura 23 Nuvem de palavras com TNIs dos investigadores. 172
Figura 24. Idade dos respondentes. 208
Figura 25. Formação dos respondentes. 208
Figura 26. Especialidade dos respondentes. 209
Figura 27. Atividades exercidas pelos respondentes. 209
Figura 28. Anos de experiência dos respondentes. 210
Figura 29. Itens utilizados pelos respondentes ao buscar artigos científicos. 211
Figura 30. Tipos de NIs na categoria Genética – Parte 1. 212
Figura 31. Tipos de NIs na categoria Genética – Parte 2. 213
Figura 32. Tipos de NIs na categoria Diagnóstico – Parte 1. 214

Figura 33. Tipos de NIs na categoria Diagnóstico – Parte 2. 214
Figura 34. Tipos de NIs na categoria Terapêutica – Parte 1. 215
Figura 35. Tipos de NIs na categoria Terapêutica – Parte 2. 216
Figura 36. Tipos de NIs na categoria Epidemiologia. 217
Figura 37. Tipos de NIs na categoria Patologia. 218
Figura 38. Tipos de NIs na categoria Imagiologia – Parte 1 219
Figura 39. Tipos de NIs na categoria Imagiologia – Parte 2 219
Figura 40. Tipos de NIs na categoria Assistência à saúde. 220
Figura 41. Tipos de NIs na categoria Pacientes. 221
Figura 42. Taxonomia dos tipos de necessidade de informação em Oncologia 223
Figura 43. Pesquisa documental sobre perfis de aplicação, esquemas e elementos de metadados. . 226
Figura 44. Pesquisa documental sobre pontos de acesso de bases de dados bibliográficas. 226
Figura 45. Atividades desenvolvidas para identificação dos tipos de necessidades de informação no domínio da Oncologia 227
Figura 46. Seleção dos tipos de necessidades de informações 228
Figura 47. Processo de sistematização dos conceitos. 238
Figura 48. As fases do DSR e a tese 244
Figura 49. Esquema lógico para a construção do perfil de aplicação. 247
Figura 50. Proposta de modelo de dados para artigos científicos no domínio da Oncologia armazenados em repositórios digitais científicos. 257
Figura 51. Interface do Language Open Vocabulary (LOV). 259
Figura 52. Modelo de dados para artigos científicos no domínio da Oncologia armazenados em repositórios digitais científicos. 298

LISTA DE TABELAS

Tabela 1. Definições para a interoperabilidade semântica 21
Tabela 2. Esquemas de codificação DC 31
Tabela 3. Identificação de perfis de aplicação de metadados 70
Tabela 4. AGRIS-AP 85
Tabela 5. Application Profile for the Americas Digital Archive 86
Tabela 6. CARLCore Metadata Application Profile 87
Tabela 7. CIHC Library Metadata Application Profile 87
Tabela 8. DC-Library Application Profile 88
Tabela 9. Dublin Core Collections Application Profile 90
Tabela 10. DPLA Metadata Application Profile 4.0 91
Tabela 11. EDM – FRBRoo Application Profile 92
Tabela 12. EThOS UKETD_DC application profile 93
Tabela 13. MPEG21 DIDL 93
Tabela 14. eSciDoc Application Profile Publication 94
Tabela 15. Europeana Semantic Elements Specification and Guidelines 95
Tabela 16. Health Canada Application Profile 96
Tabela 17. Mathematical Literature Application Profile 2010 97
Tabela 18. Michael-eu DCAP 98
Tabela 19. Mountain West Digital Library Dublin Core Application Profile 100
Tabela 20. NLM Metadata Schema 101
Tabela 21. OAI-ORE 102
Tabela 22. On Line Journal 103
Tabela 23. OSU Knowledge Bank Metadata Application Profile 103
Tabela 24. Renardus application profile 104
Tabela 25. RIOXX applicatin proifle 104
Tabela 26. RSLP Collection Description 105
Tabela 27. Scholarly Work Application Profile – SWAP 106
Tabela 28. Texas Digital Library Application Profile 107
Tabela 29. VOA3R 108
Tabela 30. OAI-ORE 108

Tabela 31. Research Data Management Application Profile 109
Tabela 32. Perfis de aplicação no domínio da Medicina 111
Tabela 33. Análise de perfis de aplicação no domínio da Medicina: elementos e propriedades dos metadados 112
Tabela 34. Análise dos perfis de aplicação no domínio da Medicina: esquemas de codificação 114
Tabela 35. Classificação dos elementos de metadados registrados nos PAMs..... 115
Tabela 36. Esquemas de codificação descritos nos PAMs 121
Tabela 37. Esquemas de metadados adotados por RDCs..... 128
Tabela 38. Bases de dados bibliográficas analisadas 129
Tabela 39. Formulário para análise das bases de dados bibliográficas..... 130
Tabela 40. SCOPUS..... 131
Tabela 41. Web of Science..... 133
Tabela 42. Annual Reviews..... 135
Tabela 43. MEDLINE 1946 to Present with Weekly Update (Ovid)..... 136
Tabela 44. MEDLINE Complete (EBSCO) 139
Tabela 45. PubMed (via National Library of Medicine) 141
Tabela 46. Bentham Science Publishers..... 143
Tabela 47. Cambridge Journals Online 144
Tabela 48. Highwire Press..... 145
Tabela 49. Science Direct..... 146
Tabela 50. Wiley Online..... 147
Tabela 51. Gale Academic OneFile 148
Tabela 52. Academic Search Premier..... 149
Tabela 53. Categorização dos pontos de acesso identificados nas bases de dados bibliográficas..... 151
Tabela 54. Necessidades informacionais: artigos identificados 168
Tabela 55. Artigos selecionados para identificação dos TNIs 173
Tabela 56. Tipos de NIs de investigadores no domínio da Medicina. 174
Tabela 57. Tipos de NIs de médicos no domínio da Medicina. 183
Tabela 58. Tipos de necessidades de informação de médicos e investigadores que atuam no domínio da Medicina..... 196
Tabela 59. População a ser investigada: Região Norte de Portugal..... 203
Tabela 60. Demonstrativo de respostas em Portugal..... 204

Tabela 61. Demonstrativo de respostas no Brasil 205
Tabela 62. Itens utilizados com maior frequência pelos respondentes ao buscar artigos científicos .. 211
Tabela 63. Tipos de NIs mais frequentes na categoria Genética 213
Tabela 64. Tipos de NIs mais frequentes na categoria Diagnóstico 215
Tabela 65. Tipos de NIs mais frequentes na categoria Terapêutica 216
Tabela 66. Tipos de NIs mais frequentes na categoria Epidemiologia 217
Tabela 67. Tipos de NIs mais frequentes na categoria Patologia 218
Tabela 68. Tipos de NIs mais frequentes na categoria Imagiologia 220
Tabela 69. Tipos de NIs mais frequentes na categoria Assistência à Saúde 221
Tabela 70. Tipos de NIs mais frequentes na categoria Pacientes 222
Tabela 71. Compilação dos resultados dos estudos desenvolvidos no âmbito da tese. 229
Tabela 72. Tipos de necessidades informacionais: definição de classes e atributos 239
Tabela 73. Identificação dos requisitos funcionais 250
Tabela 74. Definição dos *namespaces* utilizados. 261
Tabela 75. Matriz de restrições 261
Tabela 76. Esquemas de codificação de vocabulários adotados no Onco-MAP 267
Tabela 77. Esquemas de codificação de sintaxe adotados no Onco-MAP 270
Tabela 78. Questão P3. Classe: Publisher 274
Tabela 79. Questão P4. Classe: Journal 274
Tabela 80. Questão P5. Classe: Scientific Article 275
Tabela 81. Questão P6 . Classe: Author 275
Tabela 82. Questão P7. Classe: Institution 276
Tabela 83. Questão P8. Classe: Illustration 276
Tabela 84. Questão P10. Classe: Note 276
Tabela 85. Questão P11. Classe: Design Research 276
Tabela 86. Questão P12. Classe: Gene 276
Tabela 87. Questão P13. Classe: Organism 277
Tabela 88. Questão P14. Classe: Anatomy 277
Tabela 89. Questão P15. Classe: Neoplasm 277
Tabela 90. Questão P16. Classe: Disease 278
Tabela 91. Questão P17. Classe: Diagnosis 278
Tabela 92. Questão P18. Classe: Diagnostic procedure 278

Tabela 93. Questão P19. Classe: Imagiology 279
Tabela 94. Questão P20. Classe: Epidemiology 279
Tabela 95. Questão P21. Classe: Therapeutic 279
Tabela 96. Questão 2. Classe: Pharmacologic Substance 280
Tabela 97. Índice de concordância das propriedades 280
Tabela 98. Propriedades com 100% de aprovação na primeira ronda. 287
Tabela 99. Propriedades submetidas à validação na II ronda 289
Tabela 100. Classes e propriedades validadas pelo Estudo Delphi. 292
Tabela 101. Definição dos *namespaces* utilizados. 299
Tabela 102. Matriz de restrições 299

SUMÁRIO

1 **INTRODUÇÃO** 1

1.1 PROBLEMA 3

1.2 JUSTIFICATIVA 5

1.3 OBJETIVO GERAL 6

1.3.1 **Objetivos específicos** 6

1.4 ÂMBITO DA INVESTIGAÇÃO 7

1.5 PROCEDIMENTOS METODOLÓGICOS 8

1.6 ENQUADRAMENTO NA ÁREA DE SISTEMAS DE INFORMAÇÃO 8

1.7 TERMINOLOGIAS ADOTADAS NO TRABALHO 9

1.8 ESTRUTURAÇAO DA TESE 10

2 **FUNDAMENTAÇÃO TEÓRICA** 13

2.1 WEB SEMÂNTICA 13

2.1.1 ***Linked Data*** 18

2.1.2 **Interoperabilidade** 19

2.1.2.1 Interoperabilidade semântica 21

2.1.3 **Metadados** 24

2.1.3.1 Esquemas de Metadados 27

2.1.4 **Esquemas de codificação** 30

2.1.5 **Perfil de Aplicação Dublin Core** 32

2.1.5.1 Componentes do Perfil de Aplicação Dublin Core 32

2.1.5.2 Como utilizar o Perfil de Aplicação Dublin Core 33

2.2 COMUNICAÇÃO CIENTÍFICA 35

2.2.1 **Artigos científicos** 40

2.2.2 **Repositórios digitais científicos** 42

2.2.3 **Bases de dados bibliográficas** 47

2.3 NECESSIDADES DE INFORMAÇÃO 51

3 **ENQUADRAMENTO METODOLÓGICO** 59

4 **ESTUDOS PRELIMINARES** 67

4.1 PERFIL DE APLICAÇÃO DE METADADOS: PANORAMA ATUAL 67

4.1.1 **A investigação** 68

4.1.2 **O resultado** 69
4.2 *SURVEY* PARA IDENTIFICAÇÃO DE PERFIS DE APLICAÇÃO ADOTADOS POR REPOSITÓRIOS CIENTÍFICOS 122
4.2.1 **A amostra** 122
4.2.2 **O questionário** 122
4.2.3 **O resultado** 123
4.3 BASES DE DADOS BIBLIOGRÁFICAS NO DOMÍNIO DA ONCOLOGIA 128
4.3.1 **A amostra** 129
4.3.2 **A coleta de dados** 129
4.3.3 **O resultado** 150
5 ESTUDO DAS NECESSIDADES INFORMACIONAIS DOS ONCOLOGISTAS 155
5.1 NECESSIDADE INFORMACIONAL NO DOMÍNIO DA ONCOLOGIA 155
5.1.1 **Oncologia: perfil dos profissionais** 159
5.1.2 **Características das necessidades informacionais de médicos e investigadores no domínio da Oncologia** 162
5.2 *SURVEY* PARA IDENTIFICAÇÃO DOS TIPOS DE NECESSIDADES DE INFORMAÇÃO DE MÉDICOS E INVESTIGADORES NO DOMÍNIO DA ONCOLOGIA 199
5.2.1 **A amostra** 200
5.2.2 **A coleta de dados** 203
5.2.3 **O questionário** 206
5.2.4 **O resultado** 207
6 **COMPILAÇÃO DOS RESULTADOS E O CICLO DA RELEVÂNCIA DE HEVNER** 225
6.1 SELEÇÃO DOS TIPOS DE NECESSIDADES DE INFORMAÇÃO 225
6.2 TIPOS DE NECESSIDADES DE INFORMAÇÃO: IDENTIFICAÇÃO DE CLASSES E ATRIBUTOS 236
7 **DESENVOLVIMENTO DE UM PERFIL DE APLICAÇÃO NO DOMÍNIO DA ONCOLOGIA** 245
7.1 IDENTIFICAÇÃO DOS REQUISITOS 248
7.2 DESENVOLVIMENTO DO MODELO DE DOMÍNIO 252
7.2.1 **Modelo de domínio** 252
7.3 PRÉ-DESCRIPTION SET PROFILE 258
8 **ONCO-MAP: pERFIL DE APLICAÇÃO DE METADADOS** 271

8.1 VALIDAÇÃO DO PERFIL DE APLICAÇÃO DE METADADOS – ONCO-MAP 271
8.1.1 **Estudo Delphi - primeira ronda** 273
8.1.2 **Estudo Delphi – segunda ronda** 285
8.2 DESCRIPTION SET PROFILE 297
9 **CONSIDERAÇÕES FINAIS** 307
9.1 INTRODUÇÃO 307
9.2 DESENVOLVIMENTO DA INVESTIGAÇÃO 308
9.3 RESULTADO 309
9.4 AS LIMITAÇÕES E OS TRABALHOS FUTUROS 310
REFERÊNCIAS 313
GLOSSÁRIO 333
APÊNDICE A – Questionário perfil de aplicação e esquemas de metadados entre repositórios digitais científicos 339
APÊNDICE B – Cartas enviadas para instituições e profissionais para participação no inquérito sobre TNis 351
APÊNDICE C – Questionário utilizado para identificação dos TNIs médicos e investigadores 353
APÊNDICE D – Estudo Delphi – questionário primeira ronda 373
APÊNDICE E –Estudo Delphi – Questionário segunda ronda 391
APÊNDICE F – Esquema de metadados "onco" em RDF 411

1 INTRODUÇÃO[1]

> As atividades científicas e técnicas são o manancial de onde surgem os conhecimentos científicos e técnicos que se transformarão, depois de registrados, em informações científicas e técnicas. Mas, de modo inverso, essas atividades só existem, só se concretizam, mediante essas informações [...]. Fluido precioso, continuamente produzido e renovado, a informação só interessa se circula, e sobretudo, se circula livremente [...] (Le Codiac, 1996, p. 27).

A comunicação científica está associada à produção, à disseminação e ao uso de resultados de investigação científica entre os pares e para a sociedade (Garvey & Griffth, 1979). Nessa linha, Leite & Costa (2007, p. 93) entende que se trata de um "[...] conjunto de esforços, facilidades, processos dinâmicos e complexos, consensual e socialmente compartilhados, por meio dos quais o conhecimento científico – em sua vertente tácita e explícita – é criado, compartilhado e utilizado".

O processo de comunicação científica foi formalizado com a publicação dos primeiros periódicos impressos em 1665, como o *Journal de Sçavans*, fundado em Paris por Denis de Sallo, que tinha por objetivo divulgar a publicação de livros e experiências em Física, Química e Anatomia, e o *Philosophical Transactions*, da Royal Society em Londres, que focava no registro de experiências científicas (Mueller, 2000b, Swan, 2006).

Daí em diante, os avanços tecnológicos, a partir da Revolução Industrial, principalmente com o advento da Internet no fim dos anos 80, possibilitaram mudanças importantes na forma de divulgar, acessar e utilizar a literatura científica[2]. No que tange à divulgação, observa-se um número expressivo de periódicos científicos disponibilizados em formato digital, o que contribui para a rapidez da disseminação da informação. Assim como as tecnologias, as questões econômicas também impactam os periódicos. O aumento exponencial dos preços das assinaturas e os cortes no orçamento de bibliotecas inviabilizaram (e ainda inviabilizam) a manutenção ou a incorporação de novos periódicos científicos no acervo das instituições de todo o mundo, culminando com a denominada crise dos periódicos (*serial crisis*) (Mueller, 2006, Swan, 2006).

Com o surgimento, na década de 1990, da World Wide Web (WWW), houve uma aceleração na divulgação da informação científica e um aumento no acesso aos periódicos científicos. De acordo com

---

[1] Esta tese é estruturada e formatada de acordo as normas para apresentação de teses da Universidade do Minho (https://alunos.uminho.pt/PT/estudantes/Paginas/InfoUteisFormatacao.aspx) e da American Psychological Association Academy (APA). A norma da APA, elaborada para a escrita de artigo científico, orientou a padronização das figuras, tabelas, citações e referências, enquanto que, para outros aspectos, foram adotadas normas da Associação Brasileira de Normas Técnicas (ABNT): NBR-6024:2012 – Numeração progressiva das seções de um documento; NBR 6027:2012 – apresentação do sumário; NBR – 14724 - Informação e documentação: trabalhos acadêmicos.

[2] Mais informações em *Timeline* do movimento de acesso aberto: https://legacy.earlham.edu/~peters/fos/timeline.htm.

Mueller (2006) e Swan (2006), o crescimento da demanda de informações pelos investigadores e a restrição que os serviços de informação tinham em atendê-los contribuíram para o surgimento do movimento de acesso livre, que visava a quebrar o monopólio das editoras sobre os periódicos. O movimento introduziu duas modalidades de acesso livre: a Via Dourada e a Via Verde. A primeira corresponde à alternativa de possibilidade de disponibilizar o periódico científico em acesso livre. Já a Via Verde consiste em depósito de artigos científicos em Repositórios Digitais Científicos (RDCs) (Baptista, Costa, Kuramoto e Rodrigues, 2007; Harnad, 2007).

O processo de comunicação científica passa por um período de mudanças, principalmente no que se refere à disponibilização e ao acesso aos recursos, às medidas de impacto e às formas de revisão por pares. A implantação e a melhoria da qualidade dos RDCs são identificadas como grande contributo para o processo de comunicação científica. Esses repositórios se revelam uma ferramenta capaz de disponibilizar artigos científicos já revisados por pares, seja na versão *preprint* [3] ou na *post-print*, em acesso livre, criando uma alternativa ao modelo tradicional de assinaturas de periódicos estabelecido pelas editoras.

Os RDCs podem reunir, além dos artigos científicos, outros materiais digitais gerados por uma instituição – repositórios institucionais (Lynch, 2003); por uma comunidade – repositórios temáticos (Darby, Jones, Gilbert, & Lambert, 2008); ou direcionados para a educação – os que armazenam recursos de aprendizagem (Armbruster & Romary, 2010, Darby et al., 2008, Heery, 2009, Semple, 2006). Todos eles com o compromisso de armazenar, organizar e disseminar esses materiais (Darby et al., 2008, Lynch, 2003) de forma interoperável (Crow, 2002).

Um dos pré-requisitos para o sucesso dos RDCs é a interoperabilidade em níveis sintático e semântico (Patel, Koch, Doerr & Tsinaraki, 2005). A interoperabilidade semântica é definida pela European Comission como sendo o "[...] significado dos elementos de dados e os relacionamentos entre eles" (European Commission, 2010, p. 23, tradução nossa). Para assegurar que os elementos de dados possuam o mesmo significado para as partes que se comunicam, é recomendável a utilização do perfil de aplicação de metadados (PAM) (Coyle & Baker, 2009, Heery & Patel, 2000).

O PAM é uma declaração com as especificações de como os esquemas de metadados se combinam e se adaptam para descrever um determinado conjunto de recursos e explicita quais e como os elementos de metadados são adotados na descrição (National Information Standards Organization, 2007). Essa

[3] Trabalho aceito para publicação em um periódico após revisão por pares, mas que ainda não foi publicado (Packer, Santos, & Meneghini, 2017).

declaração viabiliza a utilização de linguagens (Resource Description Framework – RDF) que possibilitam a interpretação de informações por máquinas e humanos.

1.1 PROBLEMA

Segundo Diamantopoulos, Sgouropoulou, Kastrantas e Manouselis (2011), a recuperação de informações relevantes para objetivos específicos, por meio do Google Scholar, base de dados PubMed, CiteULike, e nos periódicos indexados no Directory of Open Access Journals (DOAJ), ainda requer grande esforço por parte dos investigadores durante a busca por literaturas específicas e publicações em *preprint* após arquivamento. Em alguns aspectos, os RDCs atendem a essa necessidade, mas também enfrentam problemas referentes a captura, armazenamento, organização e recuperação dos recursos, muitas vezes por não adotarem boas práticas em relação aos metadados, como a utilização e a disponibilização pública do PAM (Park & Tosaka, 2010).

Atualmente, observa-se a utilização de PAM por diversas comunidades, além de iniciativas para desenvolvê-lo, como Me4MAP, um método baseado no Dublin Core Application Profile (DCAP)[4] (Malta & Baptista, 2013).

Verificou-se até o momento que, embora existam PAMs voltados para repositórios institucionais e publicações científicas, não há um PAM específico para os RDCs que traga em seu âmbito especificações que contemplem, além dos metadados descritivos básicos (autor, título, assunto etc.), elementos de metadados que reflitam tipos de necessidades de informação (TNIs)[5] dos utilizadores em domínios específicos. No que tange a este estudo, a área em questão é a Oncologia.

A Oncologia é um domínio de alta complexidade em relação à informação e aos dados, logo, um campo de trabalho desafiador. Essa complexidade, assim como os fatores descritos a seguir, motivou a escolha dessa área para a realização deste estudo:

a) o relacionamento do câncer com o alto índice de mortalidade (8,8 milhões de pessoas morreram no ano de 2015), figurando como uma das maiores causas de morte no mundo (World Health Organization, 2016);

---

[4] Neste trabalho, são utilizados os termos Perfil de Aplicação de Metados (PAM) e Dublin Core Application Profile (DCAP) no mesmo sentido, a diferença é apenas terminólogica (cf. definições).

[5] Os termos necessidades de informação, tipos de necessidades de informação e requisitos informacionais serão definidos na Seção 2.3 Necessidade de Informação.

b) o aumento exponencial e a diversidade de informações e dados;

c) a conjugação de informações de diferentes domínios e a possibilidade de se produzir buscas em menor tempo com resultados com grau de relevância aumentado;

d) a necessidade de atualização profissional. O aumento da procura por informações pelos pacientes com câncer em todo o mundo e aspectos relacionados com o risco de morte e o crescimento do número de pacientes que possui conhecimento sobre o assunto também fazem com que os profissionais da área busquem, constantemente, atualizações (Fourie & Bakker, 2009, Stewart & Wild, 2014, World Health Organization, 2016); e

e) o reduzido número de estudos das NIs de médicos e ou investigadores que atuam no domínio da Oncologia (Fourie & Bakker, 2009).

A partir de pesquisas bibliográficas, identificou-se uma lacuna na literatura em relação a estudos sobre tipos de necessidades informacionais de investigadores e médicos no domínio da Oncologia. Essa evidência, aliada à prática profissional da autora, fez com que este estudo avançasse no sentido de identificar TNIs de investigadores e médicos que atuam em Oncologia, nomeadamente em relação ao "conhecimento médico"[6] presente em artigos científicos.

A autora acredita que, ao identificar os TNIs, seja possível traduzi-los em elementos de metadados a serem utilizados para descrição do conteúdo de artigos científicos armazenados em RDCs. Esses elementos, por sua vez, serão representados por um PAM. Ele possibilita a utilização das tecnologias da Web Semântica que permitem que sistemas de informação[7], como RDCs, bibliotecas digitais e bases de dados bibliográficas, possam ser acessados simultaneamente com a obtenção de resultados mais relevantes.

---

[6] Expressão empregada por Gorman (1995, p. 730) para distinguir os diferentes tipos de informações utilizados por médicos: dados dos pacientes (referem-se a dados pessoais, ex.: história médica); estatística populacional (agregação de dados de pacientes, ex.: dados públicos de saúde); informação logística (dados provenientes de consultorias); influências sociais (acesso à informação sobre práticas locais); e conhecimento médico (generalizável para muitas pessoas provenientes das investigações científicas, dos livros e do senso comum). O termo conhecimento médico, neste estudo, extende-se também aos investigadores.

[7] Sistemas de informação - conjuntos de componentes relacionados entre si em coleta, processamento, armazenamento, análise, disseminação e disponibilização (*web*) de informações, em determinada área, problema ou missão, com o propósito específico de atender os interesses dos utilizadores (Pinheiro & Ferrez, 2014). No contexto deste estudo, refere-se ao que é denominado de biblioteca, processamento de dados, gerenciamento da informação, suporte à tomada de decisão e sistemas de recuperação da informação.

Dessa forma, pretende-se, com este estudo, proporcionar contributos em duas frentes: a identificação de TNIs de médicos e investigadores que atuam em Oncologia e o desenvolvimento de um PAM para RDCs no domínio da Oncologia.

1.2 JUSTIFICATIVA

Observa-se que, na literatura, quando os autores se referem aos RDCs e aos artigos científicos, tópicos como acesso livre, Web Semântica[8], *Linked Data* (*Linked Open Data*) e publicações semânticas estão relacionados com a otimização do acesso e da disseminação do conhecimento. No entanto, atrelados a esses conceitos, que trazem consigo inovação, facilidades e encurtamento do caminho ao conhecimento, estão os problemas inerentes a uma era em que as mudanças acontecem rapidamente e as facilidades de disponibilização de recursos na *Web* são cada dia maiores.

Um dos grandes desafios, então, é localizar e aceder a esses mesmos recursos eficazmente, com retorno preciso e rápido. No contexto da comunicação científica, especificamente em relação aos artigos científicos e aos repositórios digitais, são identificados aspectos como o aumento exponencial de publicações *on-line*, o que alerta para a urgência de implementar ações que facilitem o acesso ao conhecimento.

A esse panorama, inclui-se a possibilidade de interligar os artigos científicos com outros recursos disponíveis em serviços de informação, como bibliotecas, arquivos e museus, com recursos do Google, como Google Maps, ou DBpedia (*Linked Data*); e explorar informações específicas, como tipo de revisão e desenhos de investigação, com a utilização das tecnologias semânticas.

Tais avanços estão associados à adoção de melhores práticas do uso de metadados para o alcance da interoperabilidade semântica, como a utilização do PAM e o uso de metadados descritivos em RDF, que podem dirimir os *gaps* existentes entre a necessidade do utilizador de identificar e aceder à informação desejada e o resultado obtido ser o esperado.

---

[8] O uso dos termos Web Semântica e *Linked Data* não é consensual. Alguns autores distinguem um termo do outro ou já veem o *Linked Data* como o termo mais atual para a ideia de Tim Berners Lee em relação à Web Semântica (Eva Mendez, 2014, informação verbal). Para Tim Berners Lee, Web Semântica é constituída de dados ligados (*Linked Data*) e *Linked Data* é a "Web Semântica feita corretamente (Bizer, Heath, & Berners-Lee, 2009, Miller, 2008). Sendo assim, o termo a ser adotado neste trabalho será *Linked Data*, por representar uma evolução da Web Semântica preconizada por Berners Lee e *Linked Open Data* (LOD) para designar dados ligados em acesso aberto.

No sentido de dirimir possíveis *gaps* e com a constatação da ausência de estudos que forneçam subsídios para a identificação dos TNIs dos utilizadores de RDC no domínio da Oncologia, este estudo estabelece a seguinte questão: quais são os TNIs de médicos e investigadores que atuam em Oncologia?

Ainda nesse sentido, a investigação está sendo desenvolvida com vista a suprir uma lacuna no contexto dos RDCs, a ausência de um PAM específico para RDC que contemple os TNIs no domínio da Oncologia, cuja realização poderá proporcionar os seguintes contributos:

a) identificação dos TNIs de médicos e investigadores que atuam no domínio da Oncologia, aspecto não encontrado na literatura até a realização do presente estudo; e

b) desenvolvimento de um PAM para RDCs com base nos TNIs dos seus utilizadores.

1.3 OBJETIVO GERAL

Desenvolver um Perfil de Aplicação Dublin Core para RDCs no domínio da Oncologia, de acordo com os princípios *Linked Open Data* (LOD) com base nos TNIs de médicos e investigadores nesse domínio.

1.3.1 Objetivos específicos

Foram definidos os seguintes objetivos específicos:

a) identificar PAMs e estudar a literatura e a documentação referente aos PAMs no domínio da Medicina, particularmente da Oncologia;

b) investigar e descrever PAMs, esquemas e elementos de metadados e esquemas de codificação utilizados por RDCs;

c) pesquisar e analisar os pontos de acesso das bases de dados bibliográficas que contemplam o domínio da Oncologia;

d) identificar e analisar, por meio de investigação documental[9] e *Survey*, os TNIs dos médicos e investigadores que atuam no domínio da Oncologia; e

e) desenvolver PAM para RDC no domínio da Oncologia utilizando o método Me4MAP.

---

[9] A pesquisa documental utiliza documentos, coleta de dados e análises como meio para obtenção de resultados consistentes. Esse tipo de investigação permite que o investigador obtenha dados que possibilitem descobrir fatos que direcionem a problemas, fontes de pesquisa, desenvolvimento de ferramentas, desenvolvimento de hipóteses ou obter conclusões (Feitosa & Feitosa, 2013).

1.4 ÂMBITO DA INVESTIGAÇÃO

A comunicação científica envolve vários tipos de materiais, entre eles os artigos científicos, que são identificados na literatura como relevantes meios para a divulgação dos resultados de investigação, independentemente de onde e de como estejam veiculados. Esses artigos podem ser disponibilizados em acesso restrito ou em acesso aberto (Via Verde ou Dourada).

Tanto os RDCs como os artigos *per se* envolvem diversos atores que buscam e produzem informação. Os produtores (editores e investigadores) de informação propiciam a divulgação, enquanto os utilizadores buscam essa informação. Entretanto, existe uma lacuna na qual se encontram dificuldades em relação ao processo de busca e recuperação: como esses materiais podem ser identificados e acessados por esses utilizadores? Quais os tipos de necessidades de informação desses utilizadores? Muitos estudos encontrados na literatura objetivam identificar causas e soluções para insatisfações decorrentes de buscas realizadas pelos utilizadores nas diversas opções de consulta (motores de busca, base de dados, bibliotecas, repositórios). Nesse sentido, faz-se necessário identificar e propiciar mecanismos que forneçam resultados com maior relevância que, no entender da autora do presente estudo, estejam associados à qualidade dos metadados, à interoperabilidade semântica, aos princípios LOD e à observância dos TNIs dos utilizadores.

Este estudo, embora busque resultados para um contexto multidisciplinar, trabalhará com um domínio de conhecimento específico (Oncologia), com um perfil distinto de utilizadores, neste caso, os médicos e investigadores; e no contexto dos artigos científicos armazenados em RDCs, a fim de restringir o âmbito da investigação.

Pelo fato de a Oncologia ser um domínio que envolve várias especialidades e níveis de formação, a inclusão dos diversos tipos de profissionais que atuam nesse domínio poderia dificultar a realização da investigação. Assim, foi priorizada a inclusão de médicos – porque eles trazem o contexto do atendimento clínico e contribuem na geração de dados para investigações – e de investigadores – pela própria característica de investigar e gerar produção científica. O contexto dos RDCs foi selecionado por ser um serviço recente, mas que tem despontado como forma eficiente de comunicação científica.

Nessa perspectiva, este trabalho analisará os PAMs, esquemas e elementos de metadados utilizados por organizações e instituições no contexto dos RDCs; publicações científicas no domínio da Medicina (Oncologia); e a estrutura das bases de dados bibliográficas (BDB). Essas análises serão realizadas na perspectiva de identificar os TNIs de investigadores e médicos que atuam em Oncologia.

Esses TNIs serão contemplados por um PAM para RDCs com o objetivo de maximizar os resultados de pesquisas dos utilizadores dos repositórios e o reuso da produção científica. Salienta-se que os benefícios decorrentes desse PAM poderão se estender aos demais recursos disponibilizados pelos RDCs, que incluem, além dos *preprints*, dados de pesquisa, monografias, teses, dissertações, *papers*, literatura cinzenta e patentes.

1.5 PROCEDIMENTOS METODOLÓGICOS

O Design Science Research (DSR) será a base para a elaboração da investigação por estar alinhado aos objetivos do estudo, à identificação dos TNIs e ao desenvolvimento de um artefacto – um perfil de aplicação de metadados. Trata-se de uma investigação com abordagem qualitativa, desenvolvida a partir da pesquisa bibliográfica, como subsídio para o embasamento teórico, e da pesquisa documental, com a obtenção de informações mais estruturadas obtidas por meio de *sites* de instituições, bases de dados bibliográficas e com consultas ao Google. O Survey foi adotado em dois estudos: na identificação de perfis de aplicação, esquemas e elementos de metadados utilizados por RDCs; e, no estudo de caso desenvolvido em Portugal, com médicos e investigadores que atuam na Oncologia. A técnica Delphi foi empregada com bibliotecários para a validação do artefacto, que foi desenvolvido com base no método Me4MAP.

1.6 ENQUADRAMENTO NA ÁREA DE SISTEMAS DE INFORMAÇÃO

À medida que a comunicação científica se utiliza das tecnologias de informação (TIs) para alcançar seu objetivo, insere-se no domínio das Tecnologias e Sistemas de Informação, que busca desenvolver e apoiar atividades para coletar, processar, armazenar, utilizar e disseminar informações (United Kingdom Academy for Information Systems, 1997). Esta investigação, que contempla duas formas de comunicação científica (artigos científicos e RDCs), relaciona-se com o domínio de Sistemas de Informação, no sentido de propiciar respostas às demandas dos utilizadores de sistemas de informação.

O domínio de Tecnologia e Sistemas de Informação propicia o desenvolvimento de ferramentas e procedimentos que auxiliam o processo de comunicação científica. No que se refere a esta investigação, pode-se citar a identificação e análise de requisitos informacionais (TNIs), a implementação de um modelo de dados, a identificação de respostas e busca de soluções para as demandas de utilizadores que usam a tecnologia como forma de acesso e recuperação da informação (Björk, 2007).

1.7 TERMINOLOGIAS ADOTADAS NO TRABALHO

Por se tratar de um estudo interdisciplinar, esta subseção visa dirimir ambiguidades em relação aos seguintes termos:

Alinhamento de vocabulário – processo de análise múltipla de vocabulários para determinar quais termos são comuns entre eles e os registros com os quais estão relacionados (W3C, 2013).

Atributo – dado significativo que pertence a uma classe, que geralmente contém valores que descrevem cada instância da classe" (Glossary of Unified Modelling..., 2017)

Categorização – é realizada por meio da análise do domínio a partir de recortes que permitem identificação dos conceitos (categorias) que fazem parte do domínio (Campos & Gomes, 2008).

Comportamento informacional – "[...] a totalidade do comportamento humano em relação ao uso de fontes e canais de informação, incluindo a busca da informação passiva ou ativa, e uso da informação" (Wilson, 2000, p. 39)

Comportamento do uso de informação – "consiste em um conjunto de atos físicos e mentais na incorporação da informação encontrada em conhecimentos existentes no indivíduo" (Wilson, 2000, p. 39).

Comportamento de busca de informação – é a "busca intencional de informação, em consequência da necessidade de satisfazer algum objetivo" (Wilson, 2000, p. 39).

Granularidade – refere-se ao tamanho, à decomponibilidade e ao grau em que um recurso é planejado para ser usado como parte de um recurso maior (NISO, 2007).

Interoperabilidade semântica – é sobre os significados dos elementos de dados e os relacionamentos entre eles. Inclui o desenvolvimento de vocabulários para descrever intercâmbio de dados e assegurar que os elementos de dados possuam o mesmo significado para as partes que se comunicam (European Commission, 2010, p. 23).

Necessidades de informação – "estado ou processo iniciado quando alguém percebe que existe uma lacuna entre a informação e o conhecimento disponíveis para resolver um problema e buscar sua solução" (Miranda & Tarapanoff, 2008, p. 1).

Ponto de acesso – nome, terminologia, código pelo qual se pode buscar e identificar um registro bibliográfico. Em sistemas de busca e armazenamento de informação por computador, campo que permite aceder a um registro ou arquivo (Young, 1988).

Precisão – é a relação entre o número de documentos pertinentes recuperados e o número total de documentos recuperados (Meadows, 1999).

Propriedade – é um aspecto específico, característica, atributo ou relação utilizada para descrever um recurso. No âmbito deste estudo, o termo propriedade foi utilizado quando se relacionava ao perfil de aplicação e o atributo, quando se referia ao modelo de domínio (ver atributo).

Recurso – um objeto abstrato que representa um objeto físico, como uma pessoa, um livro, ou um objeto conceitual, como uma cor ou a classe de coisas que têm cores.

Vocabulário – uma coleção de "termos" para um propósito específico. Os vocabulários podem variar de simples, como os utilizados amplamente pelo Resource Description Framework (RDF), Friend of a Friend (FOAF) e Dublin Core Metadata Element Set (DCMES), para os mais complexos, com milhares de termos, como os usados na área de saúde. Vocabulários desempenham um papel muito importante em *Linked Data*, especificamente para ajudar com a integração de dados (W3C, 2013).

1.8 ESTRUTURAÇAO DA TESE

Do ponto de vista estrutural, após a presente introdução, o trabalho está sistematizado em oito seções. A Seção 2, Fundamentação teórica, apresenta os tópicos Web Semântica e os subtópicos: *Linked Data*, Interoperabilidade, Metadados, Esquemas de codificação, Perfil de Aplicação Dublin Core; Comunicação científica, que engloba Artigos científicos, Bases de dados bibliográficos, Repositórios digitais científicos; e Necessidades de informação. Na Seção 3, é apresentado o Design Science Research, como metodologia a ser adotada para o desenvolvimento do estudo. Em seguida, Seção 4, são apresentados dois estudos preliminares referentes aos perfis de aplicação e um sobre às bases de dados bibliográficas. Na Seção 5, é descrita a pesquisa sobre os tipos de necessidades informacionais de médicos e investigadores que atuam na Oncologia. Na Seção 6, reúne os resultados obtidos a partir dos estudos sobre os PAMs, as Bases de Dados Bibliográficas e as TNIs de médicos e investigadores. A Seção 7, descreve o desenvolvimento de um perfil de aplicação de metadados para repositórios digitais científicos no domínio da Oncologia. A Seção 8 descreve o processo de validação do modelo de dados detalhado e define o *Description Set Profile.* Por fim, na Seção 8, apresentam-se as avaliações e os resultados esperados com a concretização do estudo. As referências consultadas estão listadas no final do trabalho e, na sequência, são apresentados os apêndices (Figura 1).

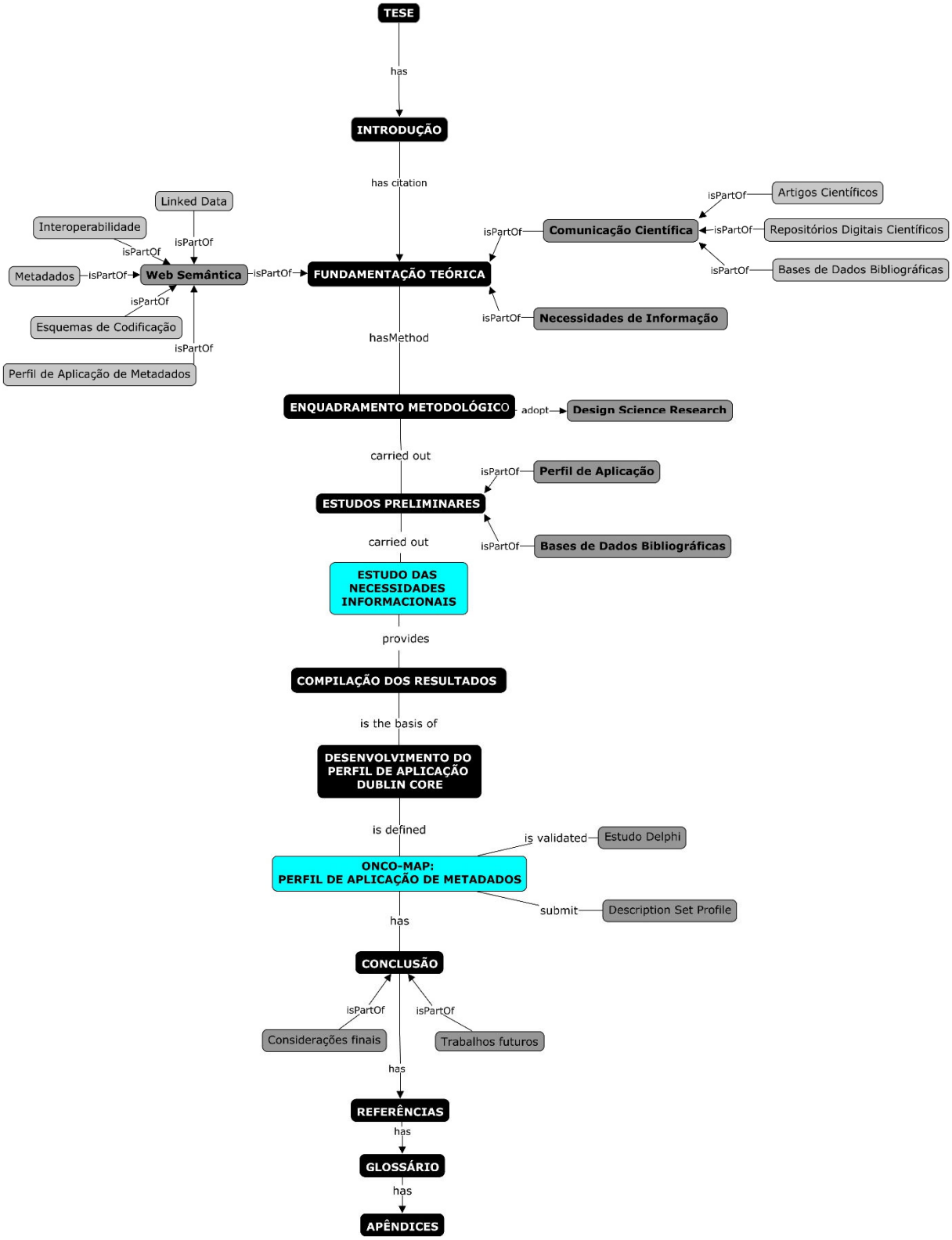


Figura 1. Estrutura da tese.

2 FUNDAMENTAÇÃO TEÓRICA

> "Cada um lê com os olhos que tem e interpreta a partir de onde os pés pisam. Todo ponto de vista é a vista de um ponto"
>
> Leonardo Boff

Neste capítulo, dissertar-se-á em relação aos tópicos necessários à elucidação da questão de investigação proposta e da consecução do objetivo deste estudo. Nesse sentido, em cada seção/subsecção, buscou-se, antes de abordar as especificidades de cada tópico, estabelecer os conceitos que norteiam o desenvolvimento da presente investigação.

Inicia-se esta fundamentação teórica pelo tópico Web Semântica, por ser o que abarca os conceitos basilares para o desenvolvimento do perfil de aplicação.

2.1 WEB SEMÂNTICA

O início da *World Wide Web* (WWW) se deu no início dos anos 90, quando Tim Berners-Lee apresentou o seu projeto de *Web* (Berners-Lee, Cailliau, Groff, & Pollermann, 1992). Uma década depois, Berners-Lee articulou a Web Semântica a partir da ideia do hipertexto concebida por Vannevar Bush, Engelbart e Ted Nelson, em que a principal característica da WWW é vincular um recurso a outro, conectando a informação de um recurso ao outro. A evolução da *Web* para uma Web Semântica possibilita o reconhecimento do significado da informação nesses documentos pelas máquinas, que pode responder se "'*head*' se refere ao chefe de uma organização ou uma coisa no topo do corpo" (Frauenfelder, 2001). Isso é possível por conta da utilização de uma semântica (significado dos elementos), de uma estrutura (organização dos elementos) e de uma sintaxe (comunicação) (Woodley, Clement, & Winn, 2013).

A sintaxe representa uma estrutura bem definida de um conjunto de símbolos mais as regras que validam a sua composição numa determinada estrutura. A semântica significa a relação entre a estrutura sintática de um símbolo e o que eles significam. Um conjunto de regras sintáticas auxilia na representação da informação, porém a ausência da semântica compromete o significado do que foi comunicado e o seu entendimento pelo homem ou pela máquina ("Semantics," 2013).

A Web Semântica, idealizada por Berners-Lee, Hendler, e Lassila (2001, p. 1) como uma extensão da *Web*, "[...] trouxe uma estrutura que propicia significado ao conteúdo das páginas *Web* [...]", criando um ambiente onde esse conteúdo é inteligível por programas especiais ou agentes inteligentes de *software*, com a possibilidade de analisar e fazer inferências sobre os conteúdos dos documentos, facilitando a

recuperação da informação pelos utilizadores. A informação é fornecida com um significado bem definido, possibilitando o trabalho cooperativo com o uso dos computadores (Berners-Lee et al., 2001, "Semantic Web", 2013).

Nessa perspectiva, a Web Semântica representa mais que uma ferramenta para administrar tarefas individuais. Adequadamente projetada, ela contribui para a evolução do conhecimento humano como um todo. Isso é possível porque a Web Semântica tem como base as quatro regras formuladas por Berners-Lee (2009):

a) primeira – presume que os recursos sejam identificados com identificador único de recurso (ou ainda Unified Resource Identified -URI). O URI é um *string* de caracteres que indica o nome ou a localização, ou ambos, de um recurso da Internet. Esse tipo de identificação utiliza protocolos específicos que permitem a conexão com representações dos recursos na *Web* ("Uniform Resource Identifier" 2013). Recentemente, faz-se referência ao Identificador de Recursos Internacionalizado ou Internationalized Resource Identifier (IRI).[10] Diferentemente de outros identificadores, como o International Standard Serial Number (ISSN), que identifica o periódico científico, o IRI informa onde ele pode ser localizado (ex.: periódico Plos One, ISSN 1932-6203, IRI http://journals.plos.org/plosone/static/publish?utm_source=dfp&utm_medium=banner&utm_campaign=one-1604-submit).

b) segunda – o uso de Hypertext Transfer Protocol (http) IRIs, utilizado para se referir a um recurso em particular;

c) terceira – deve-se fornecer informação utilizando *standards* (por exemplo, RDF, Resource Description Framework Schema - RDFS, Simple Knowledge Organization Systems - SKOS, Web Ontology Language - OWL, SPARQL Protocol and RDF Query Language - SPARQL); e

d) quarta e última regra – devem-se incluir *links* para recursos identificados por outros IRIs, de modo a descobrir mais "coisas" relacionadas.

Como se pode observar, as regras preconizadas por Berners-Lee (2009) incluem importantes tecnologias que suportam a Web Semântica, incluindo o RDF, o RDFS e a SPARQL (Figura 2). Para Castells (2003), essas tecnologias tornam a *Web* mais coesa, facilitando a localização, o compartilhamento e a integração de informações e serviços.

---

[10] The Internationalized Resource Identifier (IRI) é uma complementação para o Uniform Resource Identifier (URI). O IRI é uma sequência de caracteres do Unicode/ISO 10646 (Duerst & Suignard, 2005).

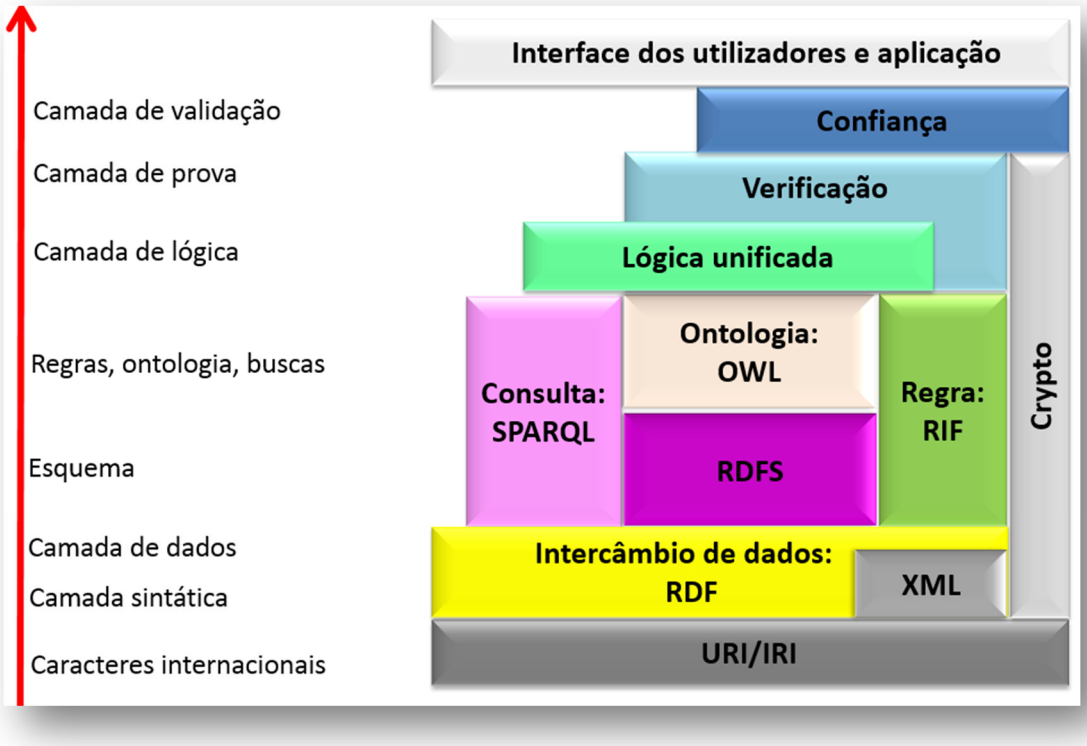


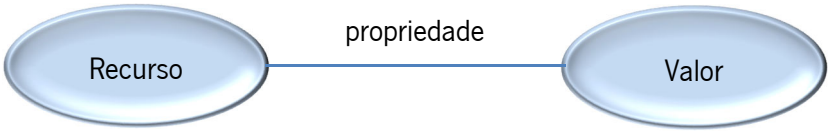
Figura 2. *Layer Cake* da Web Semântica. Adaptado de https://www.w3.org/2007/03/layerCake.png.

A partir dessas tecnologias surgiram outras que buscam extrair o maior número de dados e informações relevantes de vários recursos, como os artigos científicos, objeto deste estudo. Para entender como isso é possível, serão analisadas, a seguir, as tecnologias semânticas apresentadas no *Layer Cake* proposto inicialmente por Berners-Lee (Figura 2).

As duas primeiras camadas do *Layer Cake* – caracteres internacionais (URI/IRI) e a camada sintática (RDF/XML) – dão suporte a Web Semântica. O IRI provê a identificação única dos recursos.

O RDF "[...] é uma linguagem para representar informação sobre recursos na World Wide Web" (Manola & Miller, 2004). O RDF utiliza entre outras a linguagem eXtended Markup Language (XML) para prover a base para sua estrutura e codificação. O RDF e XML se complementam (Lassila & Swick, 1997).

O RDF fornece uma estrutura capaz de assegurar o significado das informações trocadas entre aplicações, pelo uso dos IRIs e pela forma como descreve os recursos, ou seja, descreve-os empregando declarações dos recursos ou sentenças em formas de grafos ("triplos") que consiste de nós interligados por um arco. Os nós representam os recursos e os arcos as propriedades (Manola & Miller, 2004). Os triplos podem ser representados da seguinte forma (Figura 3):



Ao serem transcritos em RDF/XML, podem ser representados da seguinte forma:

Triples of the Data Model

Number	Subject	Predicate	Object
1	http://www.aspirina.com.br/	http://www.w3.rg/2000/01/rdf-schemadrug	"Aspirina"

The original RDF/XML document

```
1: <?xml version="1.0"?>
2: <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
3:   xmlns:schema="http://www.w3.rg/2000/01/rdf-schema">
4:   <rdf:Description rdf:about="http://www.aspirina.com.br/">
5:     <schema:drug>Aspirina</schema:drug>
6:   </rdf:Description>
7: </rdf:RDF>
```

Namespace

Propriedade

Vocabulário controlado
(OBO, VIVO, RDFS, SCHEMA)

Ou graficamente:



Figura 3. Triplos RDF para descrição do recurso Aspirina.

A literatura ainda traz os termos equivalentes: sujeito para recurso, predicado para propriedade e objeto para valor (Baptista, 2002).

O RDF especifica ainda relações entre os recursos que podem ser utilizados para navegação ou integração de informações de múltiplas fontes (Baker, Bermès, Coyle, Dunsire, Isaac,.. Singer, 2011). Atende a muitas aplicações, mas é restrito em relação à descrição de propriedades, não suportando, portanto, as descrições de relacionamentos entre a propriedade e as outras fontes, daí o desenvolvimento do RDFS, que é uma linguagem para descrição de vocabulário. O RDFS também é considerado uma linguagem ontológica por permitir a especificação hierárquica de classes e propriedades necessárias para as inferências simples, embora possua algumas restrições (W3C, 2004).

Por sua vez, as ontologias, de acordo com a literatura, desempenham papel fundamental no contexto da Web Semântica. Dentre os vários conceitos de ontologias, apresenta-se o de Gruber (1993, p. 2): "[...] especificação formal, explícita de uma conceitualização compartilhada".

A ontologia define claramente o conceito adotado e os limites de seu uso. É compartilhada no momento em que reflete a noção de que capta um conhecimento consensual, não privativo de um determinado indivíduo, mas aceito por um grupo. Devido ao seu formalismo, é inteligível por computadores.

A OWL, representada no *Layer Cake* de Tim Berners-Lee, é uma linguagem (ontologia) desenvolvida como extensão do vocabulário RDF, que visa a processar o conteúdo da informação de modo a ser inteligível pelas máquinas e compreendida pelos seres humanos (Bechhofer, van Harmelen, Hendler, Horrocks, McGuinness, Patel-Schneider, & Stein, 2004). Constitui-se em um grafo RDF, ou seja, um conjunto de triplos, que pode ser escrito em diferentes formas sintáticas (RDF/XML), desde que resulte em um conjunto subjacente de triplos RDF (Antezana, Kuiper, & Mironov, 2009, Grau, Horrocks, Motik, Parsia, Patel-Schneider, & Sattler, 2008).

No processo de busca, utiliza-se a linguagem de busca (*query*) SPARQL, que é um conjunto de especificações que fornece linguagens e protocolos para consultar e manipular o conteúdo gráfico em RDF na *Web* (W3C, 2013). Desse modo, assegura-se que a informação armazenada na forma de triplos seja recuperada de repositórios com alta precisão e revocação (Antezana, Kuiper & Mironov, 2009).

Todas essas tecnologias são recomendações do W3C e, por isso, está associado a elas um estatuto de norma. Com elas e a partir da introdução de descrições explícitas do significado (semântica), os recursos disponíveis na *Web* são mais facilmente processáveis por máquinas e entendidos por humanos.

Apresentada anteriormente, a Web Semântica projetada por Tim Berneers-Lee pode ser representada como um grafo formado por nós (recursos) e arcos (relações) com identificações diferenciadas. Esses nós tipificados e interconectados utilizam classes e relações definidas por uma ontologia compartilhada entre os participantes (Castells, 2003). Na verdade, trata-se de uma base de dados gigantesca que pode ser interrogada (*queried*) utilizando mecanismos próprios (por exemplo, o SPARQL) e sobre a qual é possível aplicar mecanismos de inteligência artificial (com base em inferência) (Baptista[11], 2015, informação verbal). Para que isso aconteça, é necessário obedecer a normas, tanto relativas às

[11] Ana Alice Baptista, investigadora do Centro Algoritmi, possui doutorado em Tecnologia e Sistemas de Informação pela Universidade do Minho. É membro do Dublin Core Metadata Initiative (DCMI) Advisory Board e da International Conference on Electronic Publishing (Elpub). Atua nas seguintes áreas: comunicação científica, bibliotecas, repositórios digitais e *Linked Data*.

linguagens utilizadas (RDF, RDFS, OWL, SKOS...), aos nomes dos atributos ou propriedades (por exemplo, elementos do Dublin Core Metadata Element Set (DCMES) ou do Institute of Electrical and Electronics Engineers/Learning Object Metadata- IEEE/LOM), como aos seus tipos (por meio dos esquemas de sintaxe e de vocabulário).

Para Berners-Lee, Hendler e Lassila (2001), grande parte do conteúdo existente na *Web* não é projetada para a leitura por computadores; e, dessa maneira, o formato textual não estruturado impede que o conhecimento nele contido seja processado por programas, já que esses ainda não dispõem de um método seguro para processar a semântica. No entanto, nos últimos anos, o ambiente *Web* vem sofrendo alterações na forma como essas informações têm sido disponibilizadas. Um conjunto de diretrizes, denominado de *Linked Data*, tem possibilitado a publicação e a conexão de dados estruturados na *Web*. Para Bizer, Heath, & Berners-Lee (2009, p. 2, tradução nossa), "*Linked Data* refere-se a dados publicados na *Web* de tal forma que eles sejam inteligíveis por máquinas, o seu significado seja explicitamente definido, estejam ligados a outros conjuntos de dados externos e podem, por sua vez, estar vinculados a conjuntos de dados externos".

A adoção desse novo paradigma permite a conexão de dados de um determinado domínio com o de outros domínios, com pessoas, com outros diferentes tipos de objetos, adicionando valor aos recursos, ou seja, o *Linked Data* possibilita a publicação de informações a partir de diferentes fontes agregadas (Baker et al., 2011).

2.1.1 *Linked Data*

O conceito de Web Semântica e *Linked Data* às vezes se confundem. Para alguns autores, possuem significados distintos, mas, para Berners Lee, a Web Semântica é composta de dados ligados (*Linked Data*). Ele descreve *Linked Data* como a "Web Semântica feita corretamente". Se considerarmos esse posicionamento, a Web Semântica é o todo, e os *Linked Data*, as partes.

De acordo com Berners-Lee (2009), a diferença e a relevância da Web Semântica para a *Web* estão vinculadas à possibilidade de relacionar um dado com outro, ou com vários outros e com recursos, permitindo que pessoas ou máquinas possam explorá-los por meio dos *links* que expressem a relação entre eles (*Linked Data*).

O fato de associar um dado a outro está relacionado com a utilização do RDF, que descreve os objetos, e os IRIs que identificam o objeto ou o conceito (Berners-Lee, 2009), ou seja, *Linked Data* significa que

os dados publicados na *Web* devem estar alinhados aos quatro princípios estabelecidos por Berners-Lee, aqui já mencionados. Assim sendo, estamos a falar de *Linked Data*, que indica o uso das IRIs, sempre que possível para os valores literais e utiliza formatos e tecnologias, como RDF, OWL e SPARQL.

Nesse contexto, identificam-se os termos *Linked Data*, que corresponde à interoperabilidade técnica de dados, e *Open Data* que se refere à interoperabilidade legal (detalhes em relação à interoperabilidade serão descritos na Subsecção 2.1.2 Interoperabilidade). Quando os *Linked Data* se encontram em acesso livre, podendo ser reutilizáveis e redistribuídos, são referenciados como *Linked Open Data* (LOD) (Baker et al., 2011). A Europeana é um exemplo bem-sucedido da aplicação *Linked Data* no contexto dos serviços de informação (http://www.europeana.eu/portal/pt).

A iniciativa de colocar dados abertos em LOD é atribuída aos órgãos governamentais americanos e europeus a partir de 2009 (também seguida pelo Brasil e por outros países). Empresas de telecomunicações, como a British Broadingcasting Corporation (BBC) e o New York Times, também adotaram esse modelo ao publicar seus dados em RDF e relacioná-los com outros vocabulários da Web Semântica (Bizer, Heath & Berners-Lee, 2009). Em pouco tempo, essa mudança na publicação de dados conquistou vários adeptos, como: Bibliothèque Nationale de France; Virtual International Authority File (VIAF); Deutsche Nationalbibliothek; Europeana, com a implantação do Europeana Data Model (EDM); e a DBpedia, exemplo mais representativo dessa iniciativa (Peset, Ferrer-Sapena, & Subirats-Coll, 2011).

O ambiente onde se aplicam os *Linked Data* é constituído por uma heterogeneidade de sistemas, dados e contextos que, para se relacionarem, é preciso que sejam adotados padrões e regras que possibilitem a comunicação entre eles. A adoção de padrões técnicos, sintáticos e semânticos promove o que denominamos de interoperabilidade, a qual pode ser vista como o alicerce para a concretização do *Linked Data*, como poderá ser percebido a seguir.

2.1.2 Interoperabilidade

Miller (2000, p. 1, tradução nossa) associa a interoperabilidade às máquinas, com a preocupação de preservar a cultura organizacional e a reutilização da informação, como pode ser observado na seguinte definição: a interoperabilidade é "[...] a capacidade que os sistemas têm de se comunicarem sem requerer esforços dos seus utilizadores", ou seja, a interoperabilidade assegura que os sistemas, procedimentos e cultura de uma organização sejam gerenciados de tal forma a

potencializar o intercâmbio e o reuso da informação interna ou externamente sem intervenção humana.

A NISO (2004, p. 2, tradução nossa) conceitua interoperabilidade como "[...] a habilidade de múltiplos sistemas com diferentes *hardwares* e *softwares*, estruturas de dados e interfaces de intercambiar dados com o mínimo de perda de conteúdo e significado".

Para Patel et al. (2005, p. 6, tradução nossa), a interoperabilidade não requer um acordo preestabelecido. A interoperabilidade é identificada como "[...] qualquer forma de comunicação inter-sistemas ou a habilidade de um sistema fazer uso de dados de uma fonte não predefinida". Ou, ainda,

> [...] é a capacidade de diversas organizações interagirem em prol de benefícios mútuos e objectivos comuns acordados, que envolvem partilha de informação e conhecimento entre as organizações, por meio dos processos de negócio e do intercâmbio de dados entre os respectivos sistemas de TIC (European Comission, 2010, p. 2, tradução nossa).

Ao falar de interoperabilidade, também é necessário abordar aspectos associados aos níveis de aplicação ou compromisso, denominados por alguns autores como níveis de interoperabilidade (European Comission, 2008, Miller, 2000, Sheth, Agrawal, Lathem, Oldham, Wingate, Yadav, & Gallagher, 2006). A interoperabilidade, segundo Miller (2000), pode estar relacionada com os níveis técnico, semântico, político/humano, intercomunitário, legal e internacional.

O European Interoperability Framework (EIF) for European Public Services, que tem como proposta o desenvolvimento da interoperabilidade entre os serviços públicos europeus, abrange quatro níveis de interoperabilidade: técnico (interação e transporte), semântico (alinhamento semântico), organizacional (alinhamento entre processos e organização) e legal (alinhamento legislativo) (European Comission, 2010).

O documento preliminar para elaboração do EIF 2.0, desenvolvido pelo IDABC - European e-Government Services (European Comission, 2008), indica que a interoperabilidade em nível técnico responde às necessidades de interação entre os diversos sistemas da Tecnologia da Informação e Comunicação (TIC) no que concerne ao compartilhamento de dados por diferentes agências do governo, no contexto dos Pan-European e-Government Services, e por diferentes administrações da Comunidade Europeia. Como benefícios advindos desse nível de interoperabilidade, tem-se a redução de tempo e custo.

Em relação ao nível semântico, trata-se da possibilidade de os diferentes serviços entenderem e serem capazes de reutilizar os dados recebidos e intercambiados. Inclui acordos semânticos que viabilizam a interpretação das informações. A deficiência, nesse nível de interoperabilidade, é identificada como a

mais problemática, originando aumento do esforço e tempo necessários para a reutilização das informações.

A interoperabilidade organizacional está relacionada com as diferentes formas como os setores devem conduzir suas próprias atividades e como devem trabalhar em conjunto. A interoperabilidade política e legal proporciona aos administradores de um país, tanto de origem como de destino, o compartilhamento de dados com o mesmo significado e peso em relação à interpretação jurídica e política. Esse tipo de interoperabilidade permite que as partes interessadas possam alcançar seus objetivos legalmente (European Comission, 2008).

Assim, conforme o próprio conceito de interoperabilidade, nota-se que os níveis de interoperabilidade também não possuem consenso, mas, diante da análise das classificações adotadas por esses autores, este estudo opta pela classificação da European e-Government Services (European Comission, 2008), nomeadamente em relação à interoperabilidade semântica.

2.1.2.1 Interoperabilidade semântica

A literatura traz várias definições para a interoperabilidade semântica, como se pode identificar pelo recorte apresentado na Tabela 1. A pesquisa bibliográfica foi realizada utilizando os termos: *semantic* x *interoperable* ou *semantic interoperability* associados a contextos específicos, como *information systems*, *health*, *digital library*, *e-government* e *education*. Foram consultadas as bases de dados Scopus e Web of Science, nas quais foram contemplados os idiomas inglês, português e espanhol, limitando-se aos artigos científicos sem restrição ao período de publicação.

Tabela 1. **Definições para a interoperabilidade semântica**

Autor	Definição
Komatsoulis, Agrawal, Lathem, Oldham, Wingate, Yadav, e Gallagher (2008, p. 106)	"Semantic interoperability is the ability to access data (syntactic interoperability) and understand the data once retrieved (semantic interoperability)."
Park & Ram (2004, p. 597)	"Semantic interoperability is the knowledge-level interoperability that provides cooperating businesses with the ability to bridge semantic conflicts arising from differences in implicit meanings, perspectives, and assumptions, thus creating a semantically compatible information environment based on the agreed concepts between different business entities."
Patel et al. (2005, p. 28)	"Semantic interoperability is characterized by the capability of different information systems to communicate information consistent with the intended

Autor	Definição
	meaning of the encoded information (as intended by the creators or maintainers of the information system),"
Selvage, Wolfson, Zurek, e Kahan (2006, p. 1)	"Semantic interoperability indicates the meaning of data can be comprehended unambiguously by both humans and computer programs, and that information can be processed in a meaningful way."
Woodley, Clement, & Winn (2005)	"Ability to search for digital information across heterogeneous distributed databases whose metadata schemas have been mapped to one another. It is achieved through agreements about content description standards; for example, Dublin Core, Anglo-American Cataloging Rules" (DCMI).
He, Wang, & Liang (2010)	"Semantic interoperability depends on the system's application domain, and different application domains have different semantic interoperability objectives. In general, it refers to the capability that two *software* modules or systems can exchange the data with precise meaning, and the receiving party can accurately translate or convert the information carried by the data, including the knowledge, i.e., information and knowledge that can be understood, and ultimately produce an effective collaborative results".
Gibbons, Arzt, Burke-Beebe, Chute, Dickinson, Flewelling (2007)	"Semantic interoperability [...] has been defined as the ability of information shared by systems to be understood... so that non-numeric data can be processed by the receiving system. Semantic interoperability is a multi-level concept with the degree of semantic interoperability dependent on the level of agreement on data content terminology and the content of archetypes and templates used by the sending and receiving systems".
European Commission (2010, p. 23)	The "[...] semantic interoperability encompasses the following aspects: a) *Semantic interoperability* is about the meaning of data elements and the relationship between them. It includes developing vocabulary to describe data exchanges, and ensures that data elements are understood in the same way by communicating parties. b) *Syntactic interoperability* is about describing the exact format of the information to be exchanged in terms of grammar, format and schemas".

Ao analisar o conteúdo da Tabela 1, é possível observar que a interoperabilidade semântica definida pela maioria dos autores explicita a preocupação com a concordância no entendimento semântico. Entretanto, algumas definições são muito abrangentes e outras restritivas, em que se adotam termos que podem comprometer a compreensão do conceito de interoperabilidade semântica. Ao analisar as particularidades em termos de ambiguidade e abrangência, identificou-se que a definição apresentada pela European Comission (2010) consegue reduzir esses problemas e, portanto, será o conceito a ser adotado no contexto deste estudo.

À interoperabilidade semântica também são atribuídos níveis que tampouco compartilham de um consenso na literatura. Optou-se pela especificação dos níveis apresentada pelo DCMI, por ajustar-se melhor à proposta do estudo, como poderá ser identificado (Nilsson, Baker, & Johnston, 2009):

a) Nível 1 – definições partilhadas de termos

O Nível 1 corresponde aos termos utilizados pelo Dublin Core (DC). Esses termos são definidos com base na linguagem natural e, dessa forma, são de fácil entendimento pelas pessoas, não sendo obrigatória a utilização de IRI e estão em conformidade com os domínios e contradomínios formalmente especificados;

b) Nível 2 – interoperabilidade semântica formal

A base da interoperabilidade semântica está associada à utilização da semântica formal do RDF. Essa semântica refere-se à compatibilidade com a representação formal das declarações em RDF a partir de relações formais entre termos e regras, o que permite a realização de inferências lógicas.

Ao contrário do nível anterior, são requisitos a utilização de IRIs e a conformidade com domínios, contradomínios e relações de subpropriedades formalmente especificadas;

c) Nível 3 – interoperabilidade sintática de conjunto de descrições

Neste nível, é requisito a observação do Modelo Abstrato da DCMI (DCAM) que fornece:

- A noção de 'conjunto de descrições' como uma entidade delimitada com uma identidade específica – como na comunidade das bibliotecas e dos repositórios a noção de um "registro" que pode ser gerido.
- A noção de uma 'descrição' como uma entidade delimitada com uma identidade específica, mesmo que a descrição esteja embebida num registro.
- Um modelo estendido de 'declarações' como grafos utilizando várias combinações de URIs de Valor (Value URIs), *Strings* de Valor (Value Strings) e URIs de Esquema de Codificação de Vocabulário. (A noção formal de Esquema de Codificação de Vocabulário é específica do Modelo az da DCMI).
- Convenções para distinguir entre a 'representação' de um recurso descrito e a identificação desse recurso com um URI.
- Convenções para 'representar' um único recurso de valor com múltiplas *strings* de valor.

d) Nível 4 – interoperabilidade do perfil de conjuntos de descrições

O nível 4, representado na Figura 4, diz respeito ao PAM, que descreve as especificações de como esquemas de metadados[12] se combinam e se adaptam para descrever um determinado conjunto de recursos, explicitando quais e como os elementos são adotados na descrição dos recursos (National Information Standards Organization, 2007).

Um PAM estabelecido pela DCMI, Dublin Core Application Profile (DCAP)[13] é um documento, ou conjunto de documentos, que inclui orientação para os criadores de metadados e especificações

[12] É denominado também de vocabulários o conjunto de elementos de metadados ou de esquemas de *namespace.*

[13] Não existe distinção entre o PAM e o DCAP; é apenas uma diferença terminológica (informação verbal, Baptista, 2017).

para os desenvolvedores de metadados. O seu propósito é proporcionar a partilha de dados entre diferentes comunidades (Coyle & Baker, 2009).

Um PAM é consta de Requisitos Funcionais, um Modelo de Domínio e um Perfil de Conjuntos de Descrições cobrindo a totalidade do conjunto de metadados, como será explicitado na Subsecção 2.1.5 Perfil de Aplicação de Metadados. O Scholarly Works Application Profile (SWAP) é um exemplo desse tipo de aplicação

A Figura 4 reproduz a ideia do DCMI de que um nível é consequência do outro. Para a obtenção, por exemplo, do quarto nível, é necessário que os outros três tenham sido alcançados.

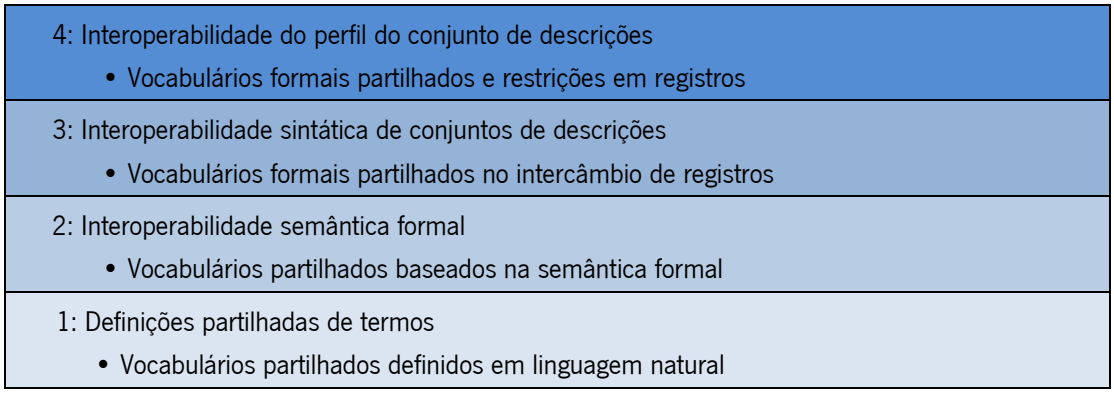


Figura 4. Níveis de interoperabilidade Dublin Core. Adaptado de: Interoperability levels for Dublin Core metadata de M. Nilsson, T. Baker, & P. Johnston (2009), retrieved from http://dublincore.org/documents/interoperability-levels/

A interpretação dada pela DCMI, em relação aos níveis de interoperabilidade semântica, será a adotada no contexto deste estudo por trazer elementos que contribuem para a realização desta investigação, como é o caso do nível 4, em que é defendida claramente a utilização do PAM. Para se entender o papel do DCAP nesse contexto, serão abordados, primeiramente, aspectos referentes aos metadados.

2.1.3 Metadados

Para a DCMI, "[...] os metadados representam a informação de um conteúdo intelectual, direitos autorais e/ou características de uma classe de um recurso de informação" (Woodley, Clement, & Winn, 2005a, tradução nossa).

Para a NISO (2004, p. 1, tradução nossa), "Metadados são informações estruturadas que descrevem, explicam, localizam e, ainda, facilitam a recuperação, o uso ou o gerenciamento de recursos de informação".

Duval, Hodgins, Sutton, e Weibel (2002, tradução nossa) definem metadados como sendo dados estruturados sobre dados. Identifica-se ainda a definição apresentada pelo *World Wide Web* Consortium (W3C), que conceitua metadados "[...] como dados sobre dados na *Web*, incluindo, mas não limitando a autoria, política, aprovação, classificação, termos de distribuição, direitos autorais, e assim por diante" ("Metadata," 2016, tradução nossa). Em relação a essas definições, observa-se que o W3C explicita o elemento *Web*, em que os metadados têm relevância significativa para a Web Semântica ("Metadata," 2016).

Vale ressaltar que a definição do W3C que relaciona os metadados com a Web retrata a tendência da disponibilização de recursos na Web, no entanto isso não quer dizer que os metadados precisam ser digitais. A Classificação Decimal de Dewey, sistema de classificação que começou a ser usado por bibliotecas em 1876, é um dos primeiros exemplos do uso de metadados. Os catálogos bibliográficos são grandes exemplos de metadados em que é possível identificar a autoria, o título e o assunto dos recursos. Na década de 1980, o padrão MARC possibilitou que os registros bibliográficos pudessem ser utilizados e compartilhados por computadores entre bibliotecas ("Metadata," 2016).

Percebe-se que, em alguns contextos e comunidades, os metadados são referidos ao dado ou à informação. Mas essa discussão recai sobre um âmbito em que não se pretende avançar – o conceito de dado e informação –, assim, procurou-se identificar apenas, dentre os conceitos citados, qual condiz mais com esta investigação. Portanto, a definição a ser empregada será a do W3C, por clarificar melhor o papel dos metadados no contexto deste estudo.

Segundo a NISO (2004, p. 3), os metadados podem ser caracterizados em razão da sua função:

a) descritivos – para identificar e recuperar recursos. Consistem de descrições padronizadas das fontes dos documentos (descrição, pontos de acesso – autor, título, editor, data de criação). Geralmente dizem respeito às informações públicas e que podem ser compartilhadas com outras instituições;

b) administrativos e de gestão – estão relacionados com a gestão de recursos. Incluem informações técnicas sobre a aquisição dos objetos digitais, o formato do seu armazenamento, proveniência, direitos autorais, licença, informações e dados necessários para a preservação; e

c) estruturais – informações sobre como colocar o objeto junto de seus componentes lógicos e facilitar a sua utilização.

Ressalta-se que, independentemente do tipo de metadado, a busca pelas melhores práticas deve ser perseguida. Com essa preocupação, a NISO (2007) formulou alguns princípios para a obtenção da qualidade dos metadados.

O Princípio 1 faz referência à necessidade de os metadados estarem em conformidade com a comunidade ou normas da comunidade, ou que permitam o seu mapeamento. Nem sempre um único esquema de metadados se ajusta às necessidades de descrição de uma coleção. É preciso utilizar elementos provenientes de diferentes esquemas, o que também acontece em relação aos vocabulários controlados, aos tesauros, aos sistemas de organização do conhecimento em geral. Por exemplo, a utilização do LOM, desenvolvido para descrição de objetos educacionais, não seria a melhor escolha para descrever publicações científicas. Nesse caso, elementos do LOM, na descrição de objetos educacionais, na descrição de objetos educacionais, podem ser utilizados juntamente com outro esquema, como o DC, para enriquecer a descrição da coleção.

As instituições devem analisar a granularidade da descrição em nível de conjunto, série ou grupo, item ou nos vários níveis, visando à profundidade desejada e ao alcance do acesso aos materiais antes de escolher um determinado esquema de metadados. É recomendado também que os esquemas adotados por instituições similares sejam considerados, pois aumentam a interoperabilidade entre coleções (tema abordado na Subsecção 2.1.2 Interoperabilidade). No entanto, a identificação dos esquemas e elementos de metadados utilizados não é o suficiente para proporcionar a interoperabilidade entre os sistemas, a recomendação é que seja implementado o PAM, que possibilita a combinação de elementos de metadados provenientes de diferentes esquemas e metadados, tópico a ser tratado na Subsecção 2.1.5 Perfil de Aplicação.

O Princípio 2 está relacionado com a interoperabilidade. "O objetivo da interoperabilidade é ajudar os usuários a encontrar e acessar objetos de informações que são distribuídos entre domínios e instituições" (p. 77, tradução nossa). Os metadados, para terem qualidade, devem ser coerentes, significativos e úteis no contexto global. Devem incluir todas as informações pertinentes sobre os objetos, como os pressupostos sobre o contexto no qual ele é acessado localmente. Os esquemas de metadados devem ser registrados de maneira a fornecer informações padronizadas para definição, identificação e utilização de cada elemento de dado. O registro de metadados define características dos metadados e os requisitos de formação, assegurando que um esquema de metadados e os elementos de dados utilizados por uma instituição possam ser aplicados de forma consistente, por ela ou pela comunidade, reutilizados por outras comunidades e interpretados pela máquina ou por humanos. O Open Metadata Registry, ou

National Science Digital Library (NSDL) Metadata Registry, como é formalmente identificado, é um exemplo de organização que mantém registros de metadados.

Quanto ao Princípio 3, os atributos dos objetos devem ser expressos utilizando termos controlados sempre que possível. Incluem-se aí nomes de pessoas, nomes de empresas, nomes de lugares, títulos e assunto. Os metadados são associados ao uso de autoridade e padrões de conteúdo para descrição dos objetos, como os sistemas de organização do conhecimento (SOCs). O Princípio 3 está relacionado também com a utilização do Vocabulary Encoding Schema (VES) (Cf. subseção 2.1.4).

Em relação ao Princípio 4, a NISO salienta que os termos e as condições de uso devem incluir informações sobre direitos autorais, se o recurso foi ou não publicado e ainda se o titular do direito autoral é identificado.

O Princípio 5 busca facilitar o gerenciamento de recursos, ou seja, são as atividades inerentes aos metadados administrativos. As informações dizem respeito à criação do objeto, quando e como, e quem é responsável pelo controle de acesso e arquivamento do conteúdo e se existe alguma restrição ao acesso.

Por fim, o Princípio 6 corresponde a metadados que fornecem informações relativas à proveniência, integridade e autoridade de um objeto, assim como a utilização de identificação única, como URI ou IRI e Digital Object Identifier (DOI).

Os metadados não são utilizados de forma isolada, são adotados esquemas. A utilização de esquemas de metadados permite aos sistemas de informação e de gestão do conhecimento a integração e o compartilhamento de informações e aplicações (Souza & Alves, 2009).

2.1.3.1 Esquemas de Metadados

Os esquemas de metadados "[...] são conjuntos de elementos de metadados concebidos para uma determinada finalidade, tais como a descrição de um determinado tipo de recurso de informação" (NISO, 2004, p. 2). Como exemplos de esquemas de metadados, podem ser citados:

a) Electronic Thesis and Dissertations Metadata Standard (ETD-MS) – padrão de metadados de teses e dissertações eletrônicas (NISO, 2004);
b) Learning Object Metadata (LOM) – metadados que descrevem objetos de aprendizagem (NISO, 2004);

c) Metadata Object Description Schema (MODS) – esquema de descrição de objetos de metadados baseado no MARC 21 (NISO, 2004);

d) Metadata Encoding and Transmission Standard (METS) – esquemas para objetos digitais de alta complexidade em bibliotecas (NISO, 2004);

e) Encoded Archival Description (EAD) – conjunto de elementos e regras para representação de partes intelectuais e físicas de inventários, guias ou catálogos (NISO, 2004);

f) Publishing Requirements for Industry Standard Metadata (PRISM) – define especificações de um conjunto de vocabulários de metadados XML para gerenciar, agregar pós-processamento de conteúdo de revistas, livros e catálogos (International Digital Enterprise Alliance, 2013);

g) Preservation Metadata: Implementation Strategies (PREMIS) – refere-se a um modelo de dados para preservação digital e Premis Data Dictionary (Caplan, 2009, Di Iorio, s.d.). Por fim, o Dublin Core que, pela sua simplicidade, tem sido o mais utilizado pela comunidade de metadados;

h) Dublin Core Metadata Element Set (DCMES ou apenas DC) – conjunto de elementos de metadados[14] Dublin Core - DC, Dublin Core Metadata Terms – termos de metadados Dublin Core. A implementação do DC buscou a interoperabilidade semântica, a flexibilidade e a modularidade na *Web*, o que levou a um consenso internacional da comunidade científica em relação à utilização desse padrão na obtenção da interoperabilidade (Pereira, Ribeiro Jr., & Neves, 2006).

A manutenção e o desenvolvimento do DC são de responsabilidade da DCMI, que busca criar mecanismos para facilitar a recuperação de recursos da Internet utilizando esquemas de metadados (Catarino, 2009).

O DC fornece um conjunto básico de elementos de descrição que podem ser adotados por catalogadores ou não catalogadores para simples descrição de recursos de informação, tornando-os mais visíveis aos motores de busca e sistemas de recuperação (Weibel, 1997). Isso é possível porque ele pode ser codificado em Hyper Text Markup Language (HTML), XML ou RDF e por ser o formato padrão empregado para efetuar a interoperabilidade entre outros sistemas. Uma abordagem mais detalhada em relação ao RDF foi apresentada na Seção 2.1 dedicada à Web Semântica.

[14] Conjuntos de elementos de metadados também são referenciados na literatura como vocabulários, esquemas de metadados ou de esquemas de *namespace.*

A descoberta de recursos por meio das comunidades de metadados é uma das questões mais importantes no ambiente digital. O DC é um dos esquemas de metadados mais utilizados e, por possibilitar flexibilidade na composição de PAMs em variados domínios, foi mais detalhado do que os demais aqui citados.

O DC compreende todos os termos mantidos pela DCMI – o conjunto de elementos DCMI Metadata Terms (DCMI Terms) e outros termos que representam propriedades, subpropriedades, classes (DCMI Type Vocabulary), esquemas de codificação Vocabulary Encoding Scheme (VES) e Syntax Encoding Scheme (SES) (Catarino, 2009). Esses elementos têm ocorrência ilimitada e são opcionais. Auxiliam na interpretação do valor de um elemento e possibilitam que utilizadores e programas interpretem esses valores

Os esquemas de metadados, assim como as relações entre os termos de metadados, são armazenados no que é denominado de registro de metadados, cuja função é fornecer descrições dos termos de metadados para humanos e máquinas (Nagamori & Sugimoto, 2006). Os registros de metadados estão relacionados com o Princípio 2 da NISO.

O esquema de registro de metadados é composto por: a) um conjunto de termos definidos para expressar propriedades de um recurso (elementos do DC, *title*, *creator* e *subject*); b) um conjunto de termos que expressa os valores de um tipo de propriedade; c) um conjunto de regras que define restrições estruturais e características neutras para alguns esquemas específicos de descrição para implementação (nível de mandato, repetibilidade); e d) um conjunto de regras vinculativas para uma XML (Heery & Patel, 2000; Nagamori & Sugimoto, 2006).

O esquema de registro de metadados é fortemente reconhecido como uma ferramenta importante não apenas para transmitir informações sobre termos para metadados, mas também para enriquecer o reuso dos vocabulários. Essa prática é essencial para que os recursos de informação sejam compartilhados e descobertos (Nagamori & Sugimoto, 2006). São exemplos de registros de metadados: Language Open Vocabulary (LOV), Open Metadata Register, BioPortal, Schema.org (health-lifesci.Schema.org), entre outros.

O acesso a esses recursos é propiciado pela qualidade dos metadados que, por sua vez, é efetivada pelo uso de esquemas de codificação.

2.1.4 Esquemas de codificação

Os esquemas de codificação do DC

> [...] apresentam informação contextual ou interpretação de regras que auxiliam na interpretação do valor dos termos. Para tanto, utiliza vocabulários controlados, notações formais ou regras de interpretação que possibilitam o entendimento por humanos, quando não forem interpretadas por um agente de software (Dublin Core Metadata Iniciative, 2012a, p. 1).

A DCMI considera dois tipos de esquemas de codificação: VES e SES. Os VES são vocabulários controlados que auxiliam na conceituação de um recurso, enquanto o SES indica que o valor é um conjunto de caracteres formatado de acordo com uma notação formal (Dublin Core Metadata Iniciative, 2012a, 2012b).

Na prática, tem-se o livro *O castelo de Guimarães*, de Barroso da Fonte, publicado em português em 10 de abril de 2000. Para descrição desse livro, utilizam-se vocabulários controlados (VES), como o Library of Congress Subject Heading (LCSH). Nesse caso, há os seguintes valores: *Castelo de Guimarães* (Guimarães, Portugal) e History - Portugal (dc.subject). Para tanto, utiliza-se o IRI: http://id.loc.gov/authorities/subjects/sh85105246 e http://id.loc.gov/authorities/subjects/sh94004918, respectivamente, permitindo que essas informações sejam legíveis por humanos e máquinas. Para que o idioma e a data de publicação também sejam legíveis para máquinas e humanos, é necessário adotar um padrão (SES). Para o idioma (português), segue-se a norma International Organization Standardization - ISO 639-2, identificada pela IRI http://purl.org/dc/terms/ISO639-2, e para o período a ISO 8601 (http://purl.org/dc/terms/W3CDTF), cujo resultado é: por e 2000-04-10, respectivamente.

O valor pode ser literal ou não literal. O literal é definido como "[...] uma entidade que utiliza uma sequência de caracteres Unicode como uma forma lexical, juntamente com uma *tag* de idioma opcional ou tipo de dados, para denotar um recurso [...]. O não literal é um valor que é físico, digital ou conceitual" (Powell, Nilsson, Naeve, & Johnston, 2007, tradução nossa). Exemplificando: para a propriedade dc:subject, tem-se o valor não literal Guimarães, Portugal (uso do LSCH); para a propriedade dc:created, é utilizado um valor literal com quatro dígitos para o ano e dois dígitos para identificar o mês e o dia, 2000-04-10. O IRI do vocabulário controlado LSCH (http://purl.org/dc/terms/LCSH) é um valor não literal.

A Tabela 2 apresenta uma síntese do que o VES e o SES representam, assim como exemplos de esquemas de codificação.

Tabela 2. **Esquemas de codificação DC**

VES	SES
Vocabulários controlados que auxiliam na conceituação de um recurso	São regras de como deve ser formatado o valor (Normas)
Sistemas de codificação: http://purl.org/dc/terms/DCMIType (dc.type) http://purl.org/dc/terms/IMT (dc.format) http://purl.org/dc/terms/LCC (dc.subject) http://purl.org/dc/terms/LCSH (dc.subject) http://purl.org/dc/terms/MESH (dc.subject) http://purl.org/dc/terms/NLM (dc.subject) http://purl.org/dc/terms/DDC (dc.subject) http://purl.org/dc/terms/UDC (dc.subject) http://purl.org/dc/terms/TGN (dc.coverage)	Esquemas de codificação: http://purl.org/dc/terms/Box (dc.coverage) http://purl.org/dc/terms/Point (dc.coverage) http://purl.org/dc/terms/ISO3166 (dc.coverage) http://purl.org/dc/terms/W3CDTF (dc.coverage) http://purl.org/dc/terms/ISO639-2 (dc.language) http://purl.org/dc/terms/RFC1766 (dc.language) http://purl.org/dc/terms/RFC3066 (dc.language) http://purl.org/dc/terms/Period (dc.date) http://purl.org/dc/terms/URI (dc.relation)
Fornece lista de valores aprovados/aceitáveis	Diz-lhe como codificar/criar/estruturar um valor, mas não enumera todos os valores possíveis
Define um valor baseado em vocabulários controlados ou sistemas de classificação	Define o valor de um campo utilizando uma notação formal
Conjunto de esquemas de conceitos em SKOS, OWL ou RDFs. Não se limita apenas a conceitos	
Os valores são não literais.	Os valores são literais.

Nota. Recuperado de: DCMI Grammatical Principles de Dublin Core Metadata Initiative, 2012a, http://dublincore.org/usage/documents/principles/, DCMI metadata terms de Dublin Core Metadata Initiative, 2012b, http://dublincore.org/documents/dcmi-terms e DCMI abstract model de A. Powell, M. Nilsson, A. Naeve, P. Johnston, 2007, *MyScienceWork*, https://www.mysciencework.com/publication/show/ad8c6acd16205d8dd583 bdeb7c8bc304.

A busca pela qualidade dos metadados e o uso do registro de metadados são fatores relevantes para a interoperabilidade semântica (NISO, 2007, W3C, 2011), assim como a utilização dos esquemas de codificação. No entanto, esse conjunto de ações, a partir do momento em que é formalizado e registrado por meio do PAM, possibilita que agentes inteligentes possam reutilizar a informação, deixando de se restringir apenas ao entendimento humano, visto que dois princípios estariam contemplados: a utilização e a explicitação de normas que estabelecem regras que proporcionam o entendimento semântico por parte das máquinas.

Com o objetivo de clarificar o papel dos PAMs no contexto deste estudo, a subsecção 2.1.5 apresenta aspectos importantes para a sua compreensão.

2.1.5 Perfil de Aplicação Dublin Core

No contexto da Web Semântica, um dos fatores que contribui para a interoperabilidade semântica é a utilização do PAM, definido por Nilsson, Baker, & Johnston (2008) como "um documento ou conjunto de documentos que especifica e descreve os metadados utilizados em uma aplicação em particular".

O PAM é aberto ao uso de diferentes esquemas de metadados desde que sejam definidos com base em RDF e sigam o modelo genérico para registros de metadados (DCAM). O PAM especifica e descreve registros de metadados que atendem a necessidades específicas de aplicação, ao mesmo tempo em que proporcionam interoperabilidade semântica com outras aplicações com base em vocabulários e modelos definidos globalmente.

Um PAM, expondo a descrição dos recursos de forma clara e objetiva, não incorrendo em ambiguidades, promove a interoperabilidade semântica a partir da utilização correta de termos definidos em um ou mais esquemas de metadados. Além da interoperabilidade semântica, o PAM também potencia a interoperabilidade sintática e estrutural (Nilsson, Baker, & Johnston, 2008).

A interoperabilidade possibilitada pelo PAM é vinculada ao modelo do RDF; ao DCAM (este caracteriza as coisas descritas pelos metadados e suas relações); ao Dublin Core Description Set Profile (DSP) (indica os termos de metadados a serem utilizados e as suas regras); e ao DCMI Guidelines for Implementation Encodings (define a sintaxe para codificar os dados utilizados pelas máquinas). Ressalta-se que o PAM promove o compartilhamento e *Linked Data* entre as comunidades (Coyle & Baker, 2009).

2.1.5.1 Componentes do Perfil de Aplicação Dublin Core

Segundo Nilsson, Baker, & Johnston (2008b), o perfil de aplicação engloba cinco componentes obrigatórios e opcionais, a saber:

a) requisitos funcionais (obrigatório) – descrevem as funcionalidades que o perfil de aplicação deve suportar e proporcionam a consistência interna e adequação do seu uso. Define o quê, como e o porquê da aplicação;

b) modelo de domínio (obrigatório) – define as entidades básicas descritas pelo perfil de aplicação e suas relações fundamentais. Refere os tipos de "coisas" (*things*) descritas pelos metadados e seus relacionamentos. Pode ser em texto ou numa linguagem, como o Unified Modelling Language (UML);

c) *description set profile* (DSP) (obrigatório) – define um conjunto de registros de metadados que são instâncias válidas de um PAM. Oferece uma linguagem simples de restrição para os metadados Dublin Core e restringe os recursos que podem ser descritos em conformidade com o perfil, assim como as propriedades que podem ser utilizadas e a forma como um valor pode ser referenciado;

d) diretrizes de uso (opcional) – orientam quanto à utilização do perfil de aplicação e das propriedades; e

e) diretrizes para sintaxe de codificação (opcional) – descrevem as sintaxes utilizadas.

2.1.5.2 Como utilizar o Perfil de Aplicação Dublin Core

A literatura traz algumas considerações em relação ao uso dos PAMs, de modo que seu objetivo seja assegurado. São elas:

a) podem ser utilizados elementos provenientes de um ou vários esquemas de metadados;

b) é vetada a criação de novos elementos a partir do perfil de aplicação – esses devem ser criados em outros ambientes e os elementos devem ser retirados de outros locais. Por exemplo, esquemas de metadados em RDFS ou em OWL;

c) havendo necessidade de criação de um *namespace*[15], o implementador deve declará-lo e mantê-lo;

d) podem recorrer a um ou mais *namespaces*;

e) especificam os esquemas utilizados e os valores, informando sobre o(s) vocabulário(s) controlados a ser(em) utilizado(s);

f) havendo necessidade, podem-se especificar padrões para elementos específicos e com formatos específicos (datas) – esquemas de sintaxe; e

g) podem refinar as definições semânticas de um *namespace*, de forma mais restrita ou particular, a fim de atender a um domínio específico ou para recursos com linguagem específica (Devey & Côté, 2006; Heery & Patel, 2000).

Embora essas informações sejam encontradas no PAM, há uma escassez de informações sobre o processo de desenvolvimento de PAMs para comunidades específicas, como pode ser constado nos estudos elaborados por Malta & Baptista (2012, 2013).

[15] Um *namespace* possibilita apropriação de um único significado para um elemento por meio da definição de estruturas, classes e funções vinculadas a ele.

Malta & Baptista (2012) desenvolveram um estudo do estado da arte em relação a metodologias para o desenvolvimento de perfis de aplicação. As autoras descreveram trabalhos relacionados com o desenvolvimento de PAMs com base na revisão de literatura e em informações obtidas diretamente de instituições, por meio de entrevistas. Com essa pesquisa, foi possível identificar manuais e artigos que contemplam as práticas de construção dos perfis, mas não metodologias formais.

As autoras verificaram que os documentos analisados foram elaborados mediante necessidades individuais e não descreveram detalhadamente o desenvolvimento dos perfis, ou são guias gerais, que também não trazem muitos detalhes. Pontuam que o desenvolvimento dos PAMs é uma tarefa difícil e, para tanto, há necessidade de maior aprofundamento no que se refere à metodologia a ser adotada para o desenvolvimento do PAM.

Dando seguimento a esse estudo, Malta & Baptista (2013) propuseram uma metodologia para o desenvolvimento de perfil de aplicação (Me4MAP) baseada no Framework for Dublin Core Application Profiles (doravante designado por enquadramento de Singapura). O modelo proposto busca suprir a ausência de guias que auxiliem na realização dos procedimentos inerentes à construção dos perfis. O Me4MAP, a ser apresentado na Secção 7, Desenvolvimento do perfil de aplicação no domínio da Oncologia, estabelece procedimentos para a construção de cada componente do enquadramento de Singapura.

SÍNTESE DA SUBSECÇÃO 2.1

Foram apresentados conceitos e aspectos pertinentes ao *Linked Data*, cujo pressuposto é a possibilidade de as máquinas poderem interpretar e recuperar as informações disponibilizadas na Internet por meio de agentes inteligentes. O *Linked Data* descreve relações entre as coisas e as propriedades das coisas. Para isso, utiliza-se de tecnologias como: RDF, RDFS, OWL e SPARQL. Essas tecnologias se baseiam em uma linguagem padronizada, cujo vocabulário é precisamente definido. Esta interpretação é possível por meio da semântica que, ao ser devidamente adotada na descrição dos recursos, proporciona a interoperabilidade semântica entre os sistemas. Com referência a esse tipo de interoperabilidade ressaltam-se as boas práticas a serem adotadas em relação ao uso de esquemas e elementos de metadados, dentre as quais se identifica o uso do PAM. Com isso, introduz-se-á na sequência o contexto para o qual será desenvolvido o perfil: os RDCs, uma das formas de comunicação científica.

2.2 COMUNICAÇÃO CIENTÍFICA

O conhecimento gerado a partir da utilização de metodologia científica, em que investigadores utilizam regras definidas e controladas, é denominado conhecimento científico (Kerlinger, 1980). Esse conhecimento é baseado em uma confiabilidade gerada pelos procedimentos científicos adotados, assim como pelo julgamento de outros cientistas. O resultado desse julgamento é feito com base no que é divulgado para a comunidade científica.

A ciência, até o século XVI, era restrita aos filósofos que utilizavam a argumentação e a dedução para explicar os fenômenos da natureza. Essa ciência era compartilhada por meio das comunicações orais realizadas em auditórios de algumas universidades que existiam na época (Björk, 2007, Mueller, 2000a). A partir do século XVII, essa dedução deixou de ser aceita como método científico, e a comunidade científica buscou outras formas de validação científica, como a observação e a experiência empírica. A partir da invenção da imprensa, foi possível proceder à sistematização da comunicação por meio dos periódicos (Björk, 2007; Mueller, 2000a). Nesse ínterim, os cientistas também começaram a organizar sociedades científicas para facilitar a divulgação do conhecimento (Björk, 2007).

O processo de divulgação e obtenção de estudos científicos é atrelado a um sistema de comunicação científica que compreende canais formais e informais. De acordo com Mueller (2000a), Borgman (2007) e Björk (2007), toda atividade de investigação produz pelo menos uma publicação formal, que pode ser vinculada a formatos impressos ou digitais: relatórios, artigos apresentados em congressos, artigos de periódicos, livros, teses e dissertações, que podem ser disseminados por meio de páginas da Internet e por repositórios digitais. Esse conjunto de publicações é denominado literatura científica. Desse leque de publicações, os artigos publicados em periódicos científicos são vistos pela comunidade científica como os mais relevantes (Mueller, 2000a).

A comunicação científica, principalmente no que se refere aos artigos científicos, permaneceu no formato impresso por muito tempo. Desde o seu início até recentemente, o formato digital era uma simples reprodução do formato impresso. A partir da década de 1990, o formato digital vem passando por evoluções que repercutem no aumento e na qualidade de acesso.

Uma das grandes transformações em relação à comunicação científica se deu a partir do movimento de acesso aberto, que teve como marco a Convenção de Santa Fé, em 1999. Esse movimento foi liderado por investigadores e acadêmicos que incentivavam a disponibilização da literatura científica em acesso livre, em resposta ao monopólio exercido pelas grandes editoras. É um movimento respaldado por três

declarações: a da Iniciativa de Acesso Aberto de Budapest, a Declaração de Bethesda e a Declaração de Berlim sobre Acesso Aberto ao Conhecimento nas Ciências e Humanidades (Baptista et al., 2007, Baptista & Ferreira, 2007, Costa, 2008, Leite, 2009, Suber, 2007).

O movimento de acesso aberto propõe dois mecanismos para a disponibilização da produção científica, denominados de vias: a Via Verde, que consiste no depósito desses trabalhos em RIs por autoarquivamento; e a Via Dourada, que é a disponibilização dessa produção dentro do conceito de acesso aberto pelas editoras (Björk, 2007, Harnad, 2007, Harnad, Brody, Vallières, Carr, Hitchcock, Gingras, ... & Hilf, 2008, Marcondes & Sayão, 2009).

Os RDCs, em que se inclui os RIs, têm se destacado nos últimos anos como potenciais meios de disseminação do conhecimento científico. Os RDCs têm como principais características o fato de reunirem, dar acesso e preservar coleção de documentos digitais; adotarem o modelo de acesso aberto e utilizarem o protocolo Open Archive Initiative Protocol for Metadata Haversting (OAI-PMH), e utilizarem o autoarquivamento (Baptista, 2010, Björk, 2007, Crow, 2002, Swan, 2011). O OAI-PMH possibilita o compartilhamento dos metadados e a interoperabilidade (Heery & Anderson, 2005). Vale ressaltar que essa prática é uma solução simples para coleta automática de metadados e possui limitações, principalmente, se analisada na perspectiva do *Linked Data*. No contexto *Linked Data*, a utilização do RDF (o uso de qualificadores torna o significado mais preciso) e de identificadores persistentes (IRI) (possibilita ligações com outros recursos) permite que informações possam ser processadas por agentes inteligentes e não apenas por humanos.

A investigação, a divulgação dos resultados da investigação por meio de publicações científicas, o acesso e a disseminação dessas publicações compõem o processo de comunicação científica analisado por autores que buscaram explicitar como esse processo ocorre (Björk, 2007, Garvey & Griffith, 1979, Hurd, 2000, Søndergaard, Andersen, & Hjørland, 2003). Björk (2007) descreve o processo de comunicação científica como um processo global, na perspectiva de "um sistema de informação conectado" (http://docs.lib.purdue.edu/cgi/viewcontent. cgi?article=1797&context=iatul)[16].

Segundo Björk, esse processo envolve criação, publicação e disseminação do conhecimento e contribui para a geração de novos conhecimentos e para a melhoria da qualidade de vida das pessoas. O autor vê os avanços tecnológicos como o agente transformador do processo de comunicação científica.

[16] Este trabalho passa a utilizar o modelo concebido por Björk sobre o processo de comunicação científica para ilustrar o contexto em que esta investigação decorre. Para tanto, nesta seção, serão utilizados apenas alguns dos seus diagramas.

Na apresentação do processo de comunicação científica, o autor se utiliza de diagramas que incluem os *inputs*, que dizem respeito ao conhecimento já existente registrado em publicações científicas, que podem ser disponibilizadas em acesso livre ou restrito, impressas ou digitais; orientações e normas que funcionam como restrições; mecanismos, como os *softwares*, que atuam no processo; e os *outputs*, que estão associados à disseminação do conhecimento.

As restrições dizem respeito aos incentivos econômicos, demandas sociais e cognitivas; às normas e aos padrões que influenciam a comunicação científica. Dentre as restrições pontuadas por Björk, identificam-se as políticas de acesso aberto, esquemas de metadados, padrão Open Archives Initiative (OAI)[17], comportamento de busca de informação[18] e as políticas de publicação.

Os mecanismos correspondem às tecnologias (*softwares*) e aos atores (informediários, utilizadores, financiadores e implementadores) que contribuem para atividades de processamento da informação, publicação, divulgação, busca, acesso e leitura, influenciadas pelo uso da Internet. Com isso o autor mostra as constribuições da academia para a construção de novos conhecimentos; e da sociedade para o aprimoramento da qualidade de vida e da produção de bens (Figura 5).

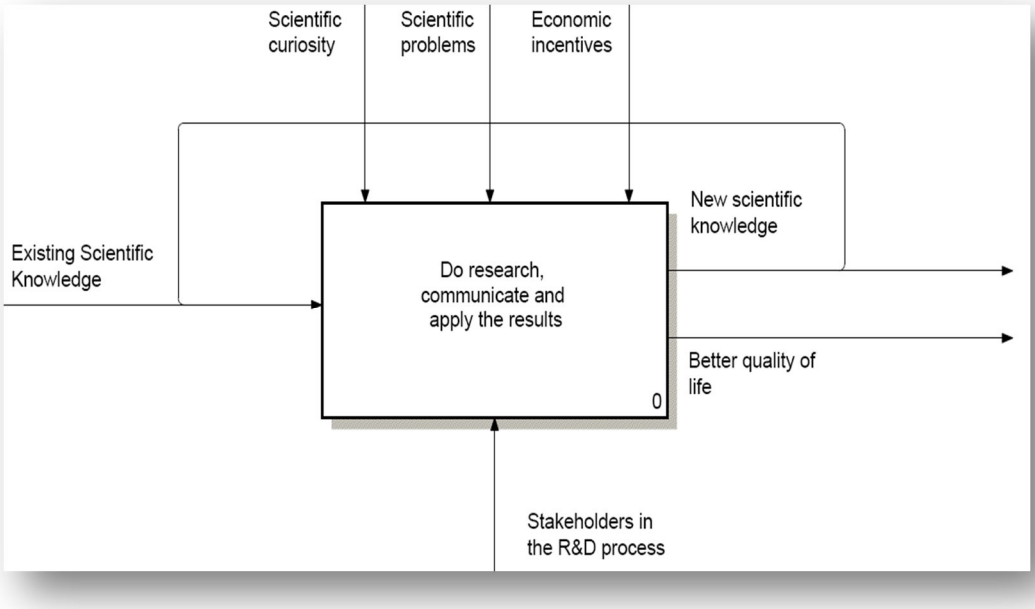


Figura 5. Fazer investigação, comunicar e aplicar os resultados - diagrama do contexto. Recuperado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 11.

---

[17] Open Archive Initiative (OAI) – desenvolveu um código em que, a partir do uso comum de determinadas *tags* (*author*, *title*, *date*), os metadados se tornam interoperáveis. Dessa forma, os metadados podem ser coletados por meio de *harvesting* e os documentos podem ser identificados e recuperados indepentemente do formato e da localização.

[18] É a "[...] busca intencional de informação, em consequência da necessidade de satisfazer algum objetivo" (Wilson, 2000, p. 39).

A Figura 6 apresenta as quatro principais etapas do processo de comunicação científica defendido por Björk. O autor destaca o papel das fontes de financiamento, com ênfase nas políticas de informação institucionais e nacionais; a investigação é apresentada como uma das atividades que exigem mais recursos; a comunicação, a mais extensa; e a disseminação do conhecimento científico representa os resultados científicos, que nem sempre são lidos pelo leitor-alvo ou, quando são lidos, nem sempre são reutilizados. A última etapa refere-se à aplicação dos resultados, que justifica os investimentos financeiros em busca da melhoria da qualidade de vida.

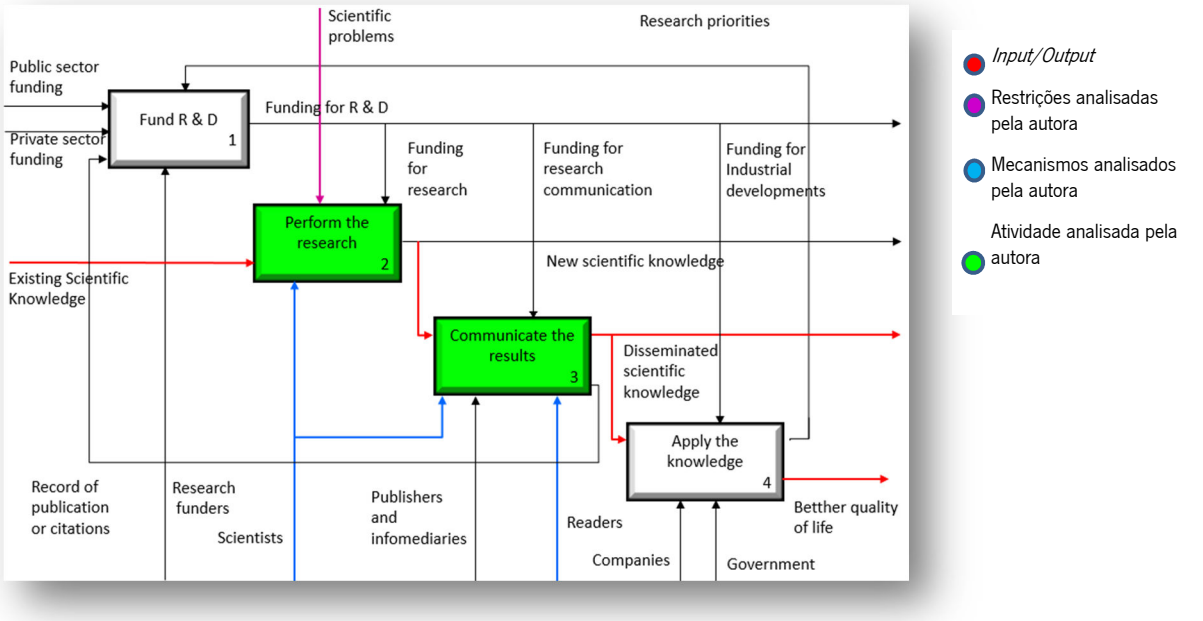


Figura 6. Fazer a pesquisa, comunicação e aplicação de resultados: desagregação (A0). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 12.

O processo de comunicação científica, a partir do advento da Internet, sofreu alterações, principalmente em relação ao formato, a exemplo das publicações eletrônicas e das redes sociais, nas quais são divulgados resultados de investigação. A Internet também produziu alterações nas atividades de captura, armazenamento, recuperação, integração e reúso de informações e dados, por conseguinte, nos papéis dos serviços e atores participantes (Björk, 2007, Shotton, 2009, Shotton, Portwin, Klyne, & Miles, 2009).

A profusão de informação que tem sido disseminada na Web, disponibilizada por bases de dados bibliográficas, repositórios digitais e por *sites*, pode levar meses para recuperação e leitura da literatura científica (Allen et al., 2003, Björk, 2007, Sandusky & Tenopir, 2008). Björk entende que é necessário empreender esforços para minimizar o tempo e o trabalho despendidos na busca e recuperação de itens

mais relevantes (Figura 7). O hábito de pesquisa de informação, abordado por Björk, corresponde às atividades desenvolvidas a partir da necessidade de atender a algum objetivo, que pode envolver o uso de sistemas de informação manuais ou informatizados (Björk, 2007, Wilson, 2000).

Para Kuruppu & Gruber (2006), é necessário compreender a necessidade informacional, comportamento de busca informacional e uso de informação[19], principalmente em relação às alterações desses hábitos, em virtude do tempo e dos avanços tecnológicos.

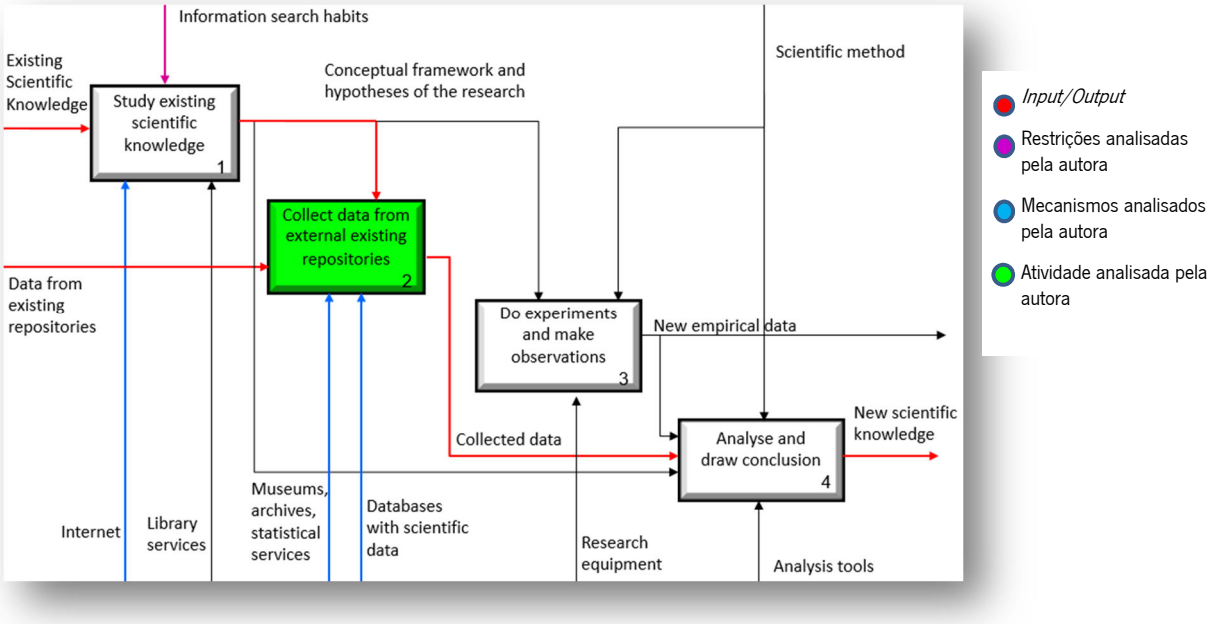


Figura 7. Performance da pesquisa (A2). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 15.

As mudanças tecnológicas emergentes implicam alterações nos suportes para as publicações eletrônicas (bases de dados bibliográficas, periódicos, RDCs e repositórios de dados). Palmer & Neumann (2002) acrescentam que a recuperação e o uso de uma determinada publicação podem ser influenciados por um tipo de estrutura de determinado domínio. A conexão estabelecida por essa estrutura termina por induzir a forma como os investigadores, principalmente aqueles de domínios interdisciplinares, localizam as publicações científicas. No caso deste estudo, os artigos científicos.

[19] Consiste de um "[...] conjunto dos atos físicos e mentais na incorporação da informação encontrada em conhecimentos existentes no indivíduo" (Wilson, 2000, p. 39).

2.2.1 Artigos científicos

A comunicação científica tem sido realizada em forma de teses e dissertações, livros, *papers* apresentados em eventos científicos, sendo uma das mais consagradas os artigos publicados em periódicos científicos. O *Journal de Sçavans*, fundado em Paris por Denis de Sallo, em 1665, é referenciado como o primeiro periódico científico. Teve por objetivo divulgar a publicação de livros e experiências em Física, Química e Anatomia. Já o periódico *Philosophical Transactions*, publicado no mesmo ano pela Royal Society, em Londres, tinha como objetivo apenas o registro de experiências científicas.

Os periódicos científicos possuem o objetivo de controlar e certificar a qualidade dos resultados de investigação, divulgar o conhecimento, atualizar os investigadores em relação aos avanços científicos e proteger legalmente os direitos do autor (Schaffner, 1994).

Björk (2007) caracteriza os periódicos como publicações com uma definição clara; que priorizam novas descobertas cujo conteúdo é constituído de textos escritos. Também podem ser disponibilizados em acesso restrito ou aberto após serem avaliados por pares (*peer-review*).

O processo de revisão por pares teve início no século XVII e corresponde à avaliação crítica de manuscritos de investigações realizada por especialistas da área do estudo. Existem dois tipos de sistemas, o *open review* (sistema aberto) e o *double blind review* (sistema duplo cego) (Jenal, Vituri, Ezaías, Silva, & Caliri, 2012).

Cada periódico tem um perfil que é direcionado a uma comunidade específica de leitores que influencia a estrutura dos artigos, como tamanho do artigo, número de referências, processo de revisão por pares e estrutura de citações (Schaffert, Gruber, & Westenthaler, 2005). Os periódicos possuem uma identificação única (ISSN) e os artigos também podem ser identificados univocamente por meio do DOI.

A disponibilização dos periódicos científicos em formato eletrônico não proporcionou apenas facilidades de acesso aos artigos científicos digitais, trouxe também desafios relacionados com a recuperação, independentemente do sistema utilizado para armazenamento: bibliotecas digitais, bases de dados bibliográficas ou RDCs. Esses desafios estão associados em parte ao crescimento exponencial de informação (na Medline são incluídos em média 1.300 artigos por dia[20]), à impossibilidade de o ser

[20] https://www.w3.org/wiki/HCLSIG/SWANSIOC/Actions/RhetoricalStructure.

humano processar grandes volumes de informação e à dificuldade que as máquinas têm de identificar e interpretar o conteúdo semântico desses artigos.

Por meio de tecnologias que exploram a semântica e a agregação de dados, os artigos estão se tornando mais dinâmicos e descritivos, como pode ser conferido nos seguintes projetos: Project Prospect and the Royal Society of Chemistry, The ChemSpider Journal of Chemistry, PubMed Central and Biolit, Public Library of Science (PLoS) (Attwood, Kell, McDermott, Marsh, Pettifer, & Thorne, 2009, Groth, Gibson, & Velterop, 2010, Shotton et al., 2009).

No entanto, o que se busca neste estudo é a utilização das tecnologias semânticas aliadas aos princípios *Linked Data* como potencializadoras para a descoberta de artigos científicos, no caso particular desta investigação, armazenados em RDCs (Figura 8). Uma vez que os elementos de metadados desses artigos são publicados na Web, é necessário que algumas práticas sejam adotadas (elementos de metadados interoperáveis, definidos em esquemas de *namespaces* com esquemas de codificação apropriados) para que buscas sejam realizadas simultaneamente em sistemas de informação distintos e ainda assim propiciem resultados com maior índice de relevância.

As consultas por artigos em grandes coleções são realizadas utilizando palavras-chave, ações de robôs e/ou operadores boleanos em que as informações recuperadas, como já foi comentado, nem sempre refletem à necessidade do utilizador. Mas para Berners-Lee et al. (2001, p. 10), apesar dos "*The computer doesn't truly 'understand' any of this information* [...] *it can now manipulate the terms much more effectively in ways that are useful and meaningful to the human user*" (Web Semântica).

Outros autores também veem no *Linked Data* a possibilidade de reduzir a ambiguidade por meio de identificação de conceitos e conexão de recursos com base na semântica (Ambinder & Marcondes, 2012, Mons & Velterop, 2009).

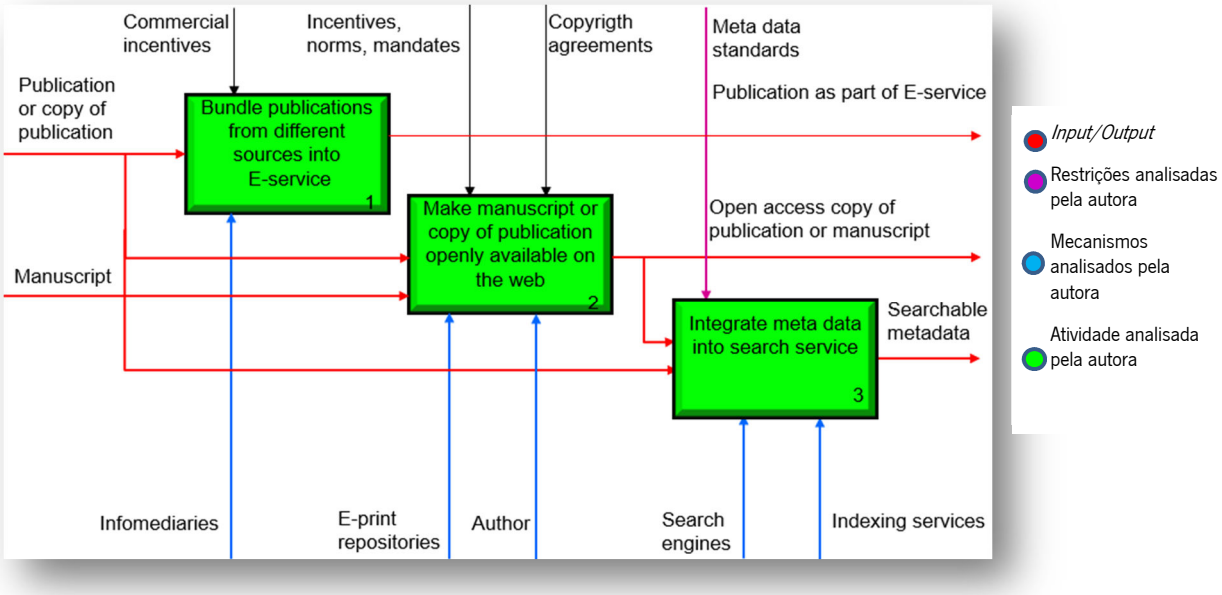


Figura 8. Facilitar a recuperação global (A3221). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 28.

Como descrito, o advento da tecnologia alterou sobremaneira a forma como os resultados de investigações científicas são divulgados. Dentre as alterações identificadas na literatura, destaca-se a possibilidade de disseminar os artigos científicos em acesso aberto. Como referido anteriormente, são duas modalidades de acesso livre: a Via Dourada, está associada à disponibilização de periódicos científicos em acesso livre, e a Via Verde permite que artigos científicos sejam disponibilizados por meio dos repositórios digitais científicos, tópico a ser apresentado na próxima subseção.

2.2.2 Repositórios digitais científicos

Um repositório digital é um "[...] sistema em rede que fornece serviços pertencentes a uma coleção de objetos digitais" (Bekaert & Van De Sompel, 2006).

A implantação dos RDCs foi impulsionada pela utilização dos protocolos de coleta de metadados OAI-PMH, desenvolvidos com o intuito de melhorar a capacidade de descobrir documentos acadêmicos e dados disponibilizados na Internet, fornecendo uma estrutura de interoperabilidade aplicativo-independente, baseada na recolha/coleta de metadados. Esse protocolo é um conjunto de seis verbos[21] ou serviços que invocam o http, integra ainda os fornecedores de dados (repositórios), administradores

[21] Identify, ListIdentifier, ListMetadataFormats, ListSets, ListRecord, GetRecord.

de sistemas, que expõem a estrutura dos metadados via OAI-PMH, e os fornecedores de serviços, que efetuam a recolha/coleta desses dados como base para a construção de um serviço de valor agregado (Lagoze, Van de Sompel, Nelson, & Warner, 2002).

Os RDCs participam de forma impactante para a disseminação do conhecimento científico em vista dos seguintes aspectos:

a) publicações de diferentes fontes em serviços de informação informatizados, em que se inclui o processo de autoarquivamento;

b) disponibilização de manuscritos (*preprint*) em acesso livre na Web. Para tanto é necessária a institucionalização de políticas de informação que asseguram o acesso livre à informação (mandatória); a adoção de Licenças Creative Commons; e

c) integração de metadados por mecanismos de busca. Deve-se considerar a qualidade de metadados, que promove a interoperabilidade semântica e a aplicação dos princípios *Linked Data*; e a representação dos documentos por meio de elementos de metadados (Baptista, 2010, Björk, 2007, Crow, 2002, Lagoze et al., 2002, Park & Tosaka, 2010, Swan, 2011) considerando, por exemplo, os TNIs dos utilizadores (Figura 9).

Os repositórios podem ser classificados em institucionais (RI), temáticos (RT) e de objetos de aprendizagem (Armbruster & Romary, 2010, Darby et al., 2008, Heery, 2009, Semple, 2006). O RI foi definido por (Lynch, 2003, p. 1, tradução nossa) como "[...] um conjunto de serviços que uma universidade oferece aos membros de sua comunidade para o gerenciamento e a disseminação dos materiais digitais gerados pela instituição e por seus membros", cujas características são: ser digital, institucionalmente definido, acadêmico, cumulativo e perpétuo, com acesso livre e interoperável (Crow, 2002).

Os RTs, como o arXiv ou Cogprints, possuem conteúdo voltado a um domínio específico, e os colaboradores podem ser de instituições distintas, provenientes de diferentes países (Darby et al., 2008). O arXiv possui como particularidade o fato de ter sido criado em 1991, ainda quando não existiam periódicos eletrônicos, com o propósito de possibilitar o acesso global a *preprints* de artigos científicos no domínio da Física.

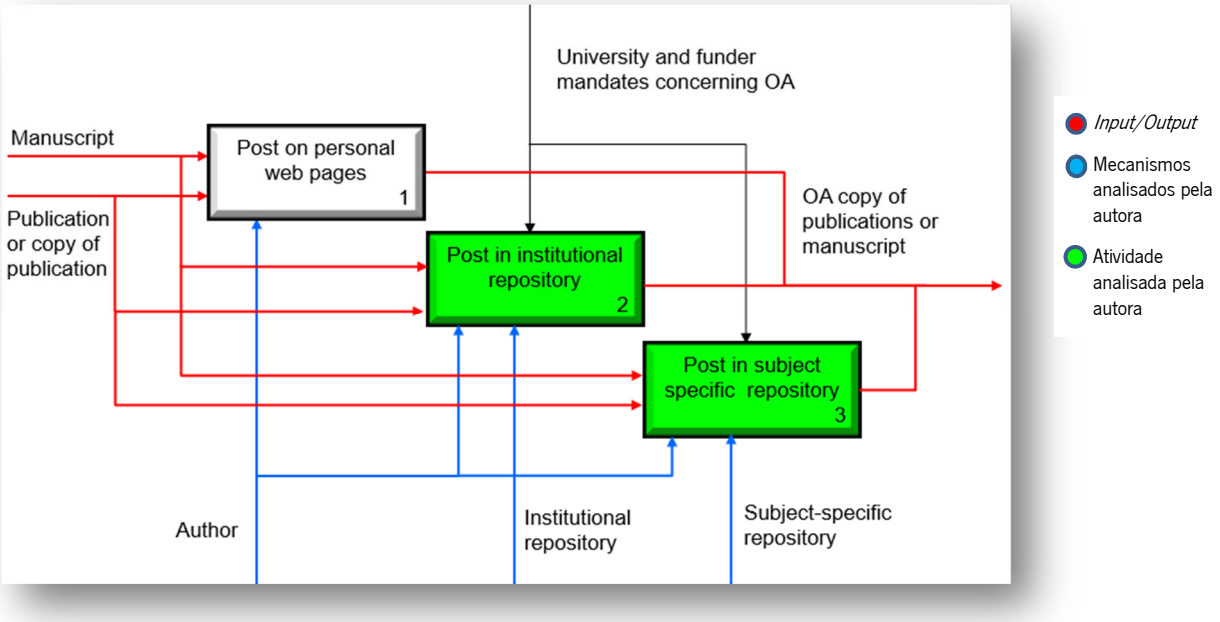


Figura 9. Fazer manuscrito ou cópia da publicação em acesso aberto (A32212). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 29.

Ao analisar estudos em relação às vantagens dos repositórios, identifica-se que eles vão além da academia, beneficiam a sociedade como um todo ao disponibilizar publicações científicas sem as restrições impostas por editoras comerciais (Drake, 2004, Foulonneau, Cole, Habing, & Shreeves, 2005, Lynch, 2003). Muitos esforços têm sido empreendidos no sentido de melhorar a qualidade dos repositórios, e a busca pela interoperabilidade semântica entre os RDCs se apresenta como um dos maiores desafios (Baptista, 2010, Bhat, 2010, Park & Tosaka, 2010, Swan, 2011).

O diagrama que ilustra a fase de integração dos metadados traz uma etapa importante para que a recuperação dos recursos possa proporcionar resultados relevantes para os utilizadores (Figura 10). Fazem parte desse diagrama os serviços de indexação, que atualmente vão além dos serviços realizados por profissionais da informação ou por sistemas inteligentes. Os utilizadores também estão a executar atividades que há alguns anos eram inerentes aos profissionais de informação, como o autodepósito que inclui a associação de palavras-chave ou a inclusão de *tags* (folksonomias) para descrição do recurso.

Alguns procedimentos são exigidos de forma a viabilizar a recuperação dos recursos independentemente do serviço de informação. Esses procedimentos estão relacionados com o uso de IRIs, as boas práticas de uso de metadados, a interoperabilidade semântica e o OAI-PMH. O OAI-PMH é uma aplicação simples que possibilita a coleta de metadados em diferentes repositórios reunidos em um único local. Quando um repositório é criado, um mínimo de elementos de metadados Dublin Core é disponibilizado,

possibilitando uma base mínima para o compartilhamento dos recursos, ou seja, fornece um nível básico de interoperabilidade (Open Archives Initiative Dublin Core -OAI-DC) (Tedd & Large, 2004).

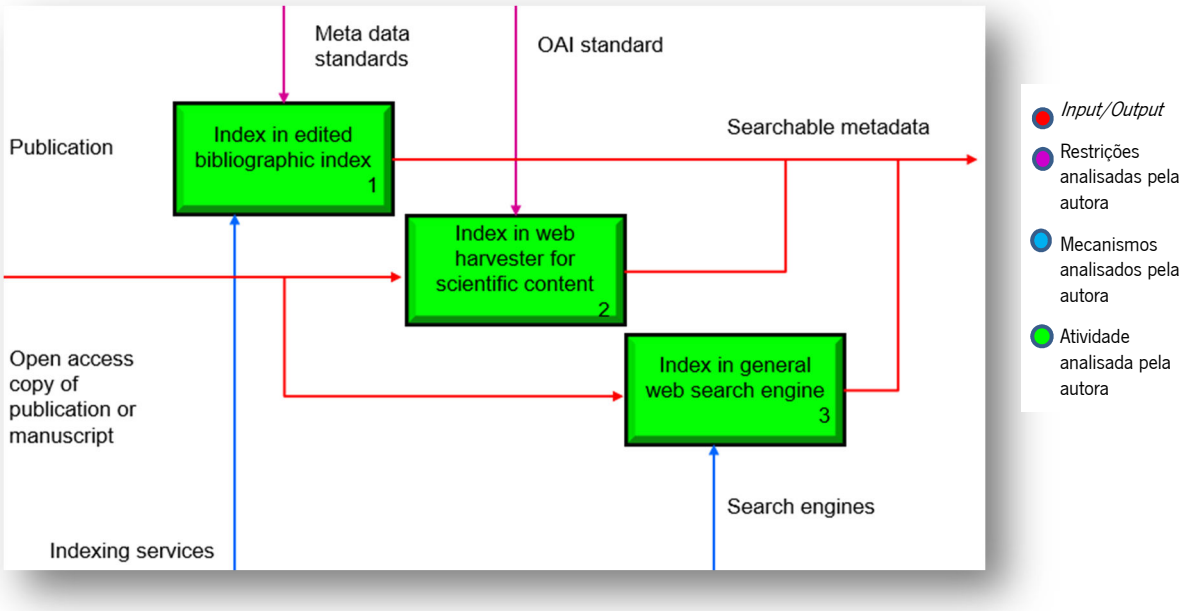


Figura 10. Integração de metadados nos serviços (A32213). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 30.

Ainda em relação à indexação dos recursos, é importante o papel que as tecnologias semânticas têm nesse processo, a começar pela adoção do SKOS que dá suporte ao compartilhamento e ligação com os sistemas de organização do conhecimento (SOCs ou KOS, em inglês), como vocabulários controlados, tesauro, ontologias e esquemas de classificações na *Web*.

Conforme abordado, as atividades de busca de informação são norteadas pelo comportamento informacional[22] dos utilizadores principalmente em relação à Web, motores de buscas, serviços de informação e consultas ao acervo de bibliotecas. A identificação dos TNIs pode contribuir para uma recuperação mais eficiente e eficaz à medida que traz para os sistemas especificidades que dizem respeito ao domínio, contexto e atividades do utilizador (Figura 11).

[22] "[...] a totalidade do comportamento humano em relação ao uso de fontes e canais de informação, incluindo a busca da informação passiva ou ativa, e uso da informação" (Wilson, 2000, p. 39).

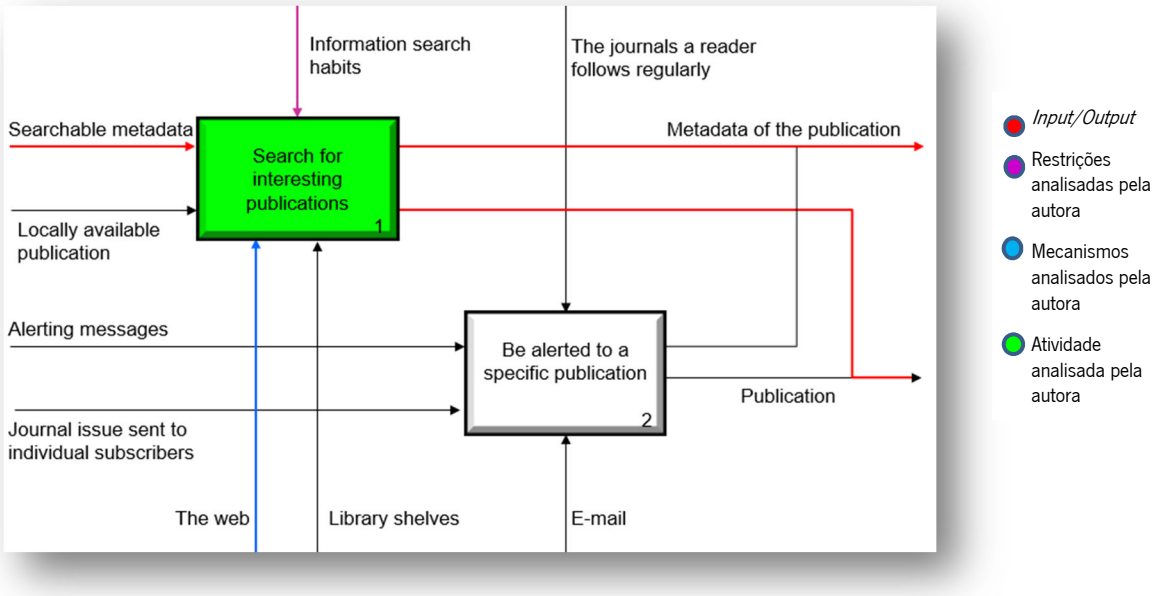


Figura 11. Conheça a publicação (A3231). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 33.

Recordando que não se está a falar apenas da *Web* dos documentos, dos hipertextos, as iniciativas a serem adotadas visam a atender à busca por informação em um ambiente diversificadamente rico e, portanto, a possibilidade de conectá-los traz inúmeras formas de acesso e reúso. A exploração desses recursos é significativamente ampliada. Nesse sentido, a identificação e o aproveitamento dos TNIs dos utilizadores na descrição dos recursos podem ser vistos como uma contribuição inestimável para investigadores, desenvolvedores, provedores e utilizadores de unidades de informação. No entanto, vê-se também aqui um desafio, que tem a atenção de profissionais da informação há décadas, e que deve ser impulsionado, já que esses TNIs podem ser integrados a um ambiente semântico. Portanto, normas e padrões devem ser seguidas a fim de proporcionar a interoperabilidade semântica e sintática entre sistemas.

Os esquemas de metadados e o padrão OAI apontados por Björk sugerem que a descrição e a recuperação dos recursos sejam realizadas em um contexto que pressupõe a existência de interoperabilidade, em que os sistemas podem "conversar", otimizando a recuperação dos recursos. Nesse caso, o PAM, que tem o papel de intermediar esse diálogo, também poderia ser incorporado à estrutura do modelo como restrição (Figura 10).

Ainda nesse sentido, as linguagens RDF e OWL integrariam o diagrama como restrições capazes de potencializar a descrição e o acesso aos recursos. Logo, pode-se fazer um paralelo desse diagrama com as atuais formas de descrição dos recursos na *Web*, em que a utilização de tecnologias semânticas,

como as linguagens RDF, RDFS e OWL, possibilita a ligação entre diversos tipos de recursos e dados armazenados em sistemas de informação distintos – *Linked Data*. Assim, nesse modelo, ainda se pode utilizar a linguagem SPARQL que se reflete na qualidade de recuperação da informação armazenada em RDF.

Sobre o atual panorama dos RDCs, de acordo com o Registry Open Access Repositories (http://roar.eprints.org/), dos 3.430 repositórios identificados no *site*, 2.990 armazenam publicações científicas e apenas 19 informam que seguem as diretrizes do LOD. Vale ressaltar que a baixa quantidade, baseada nos dados disponibilizados pelo Diretório, foi confirmada pelos resultados de investigação sobre o uso de PAMs por repositórios (Andrade & Baptista, 2015).

Esses resultados merecem uma análise atualizada em relação à adoção das diretrizes *Linked Open Data* pelos RDCs, como foi realizada para museus, bibliotecas e arquivos (https://journals.ala.org/ltr/article/view/ 5892/7445). São necessárias ações que assegurem a qualidade desses repositórios, principalmente em relação ao *Linked Data*. Ou seja, a implementação dessas ferramentas nesse contexto vai ao encontro do que Björk invoca a implementação de iniciativas que proporcionem aos utilizadores o acesso ao conhecimento em menor tempo e com esforço reduzido (Figura 11).

A partir dessas considerações, a autora sugere a inclusão das ferramentas semânticas e a integração dos princípios *Linked Open Data* ao processo proposto por Björk, no que se refere à descrição, à representação e à recuperação dos documentos.

2.2.3 Bases de dados bibliográficas

Uma base de dados bibliográfica corresponde a uma coleção organizada de recursos bibliográficos ou não bibliográficos/dados em formato legível por computadores. A maioria delas contém registros descritivos de livros, patentes, artigos científicos e anais de conferência e representa um dos melhores exemplos de armazenamento e disseminação da informação científica no contexto da Web. Reúne e organiza referências (bases referenciais) e/ou conteúdos (bases de texto completo) com a utilização de estruturas de dados que possibilitam pesquisas por meio de campos de busca (pontos de acesso), o que contribui para a recuperação dos artigos científicos (“Bibliographic database - New World Encyclopedia,” s.d.). Para Björk, as bases de dados bibliográficas são consideradas como a mais tradicional fonte de pesquisa estruturada.

As buscas podem ser realizadas por títulos, nomes de autores, afiliação, resumos, palavras-chave ou utilizando sistemas de organização do conhecimento, como tesauros, taxonomias, vocabulário controlado, ontologias (GoPubMed - www.gopubmed.org/goweb/) e identificadores persistentes (IDs), como o ID da PubMed e o DOI.

As BDBs surgiram nos meados dos anos 60 e a Medical Literature Analysis and Retrieval System (MEDLARS) foi uma das precursoras. A Chemical Abstracts Services (CAS), fundada pela National Science Foundation, the National Institutes of Health, and the Department of Defense, foi a primeira base de dados a investir no desenvolvimento e investigação na implantação de bases de dados bibliográficas. Em 1965 foram disponibilizadas a CAS e a Chemical and Biological Activities (CBAC), no formato impresso e magnético. Outras bases foram implementadas a partir de 1969, no domínio de Engenharia, a Engineering Index e a BioSciences Information Services (BIOSIS) (Neufeld & Cornog, 1986). Desde então, várias bases vêm sendo criadas e disponibilizadas ("Bibliographic database - New World Encyclopedia," n.d.), a exemplo da Medline, que iniciou suas atividades em 1971.

A Medical Literature Analysis and Retrieval System Online (Medline) é uma base de dados referenciais no domínio das Ciências da Saúde de livre acesso, gerenciada pela National Library of Medicine, Estados Unidos (EUA). Ainda no domínio da Medicina, tem-se a Excerpta Medica Database (EMBASE) e a UptoDate. Esta última é considerada uma das mais importantes bases de dados em Medicina Baseada em Evidências. Ambas são de acesso restrito. As BDBs ainda podem contemplar coleções de uma editora específica, como a ScienceDirect, produzida pela Elsevier e ainda podem ser divididas em multidisciplinar (ProQuest, Web of Knowledge, Scopus) e especializada (Anthropological Index Online, Education Resources information Center - ERIC).

Existem BDBs de livre acesso e outras que requerem inscrição. Muitas instituições pagam grandes valores por essas assinaturas (Ovid, EBSCO e Web of Science). No contexto do acesso aberto, a Medline, os *links* podem direcionar para periódicos com *peer-review* ou com controle de qualidade em acesso livre no *site* da PubMed Central.

Na Figura 12, Björk apresenta o uso da informação por algumas categorias de usuários, em que os médicos procuram se manter atualizados e os acadêmicos (investigadores) buscam desenvolver investigações. Embora as BDBs sejam o serviço de informação mais tradicional, o acesso livre à informação é relevante para que esses profissionais alcancem seus objetivos e continuem a produzir novos conhecimentos ou apliquem em ações que melhorem a qualidade de vida (Figura 13), daí a relevância de prover avanços em relação aos RDCs. Para tanto, é necessário promover mudanças

fundamentais na comunicação científica; adaptar-se aos novos paradigmas, em que novas tecnologias estão a emergir de maneira a explorar a semântica dos dados, para que sistemas inteligentes sejam capazes de atender às NIs dos utilizadores com maior eficácia e eficiência.

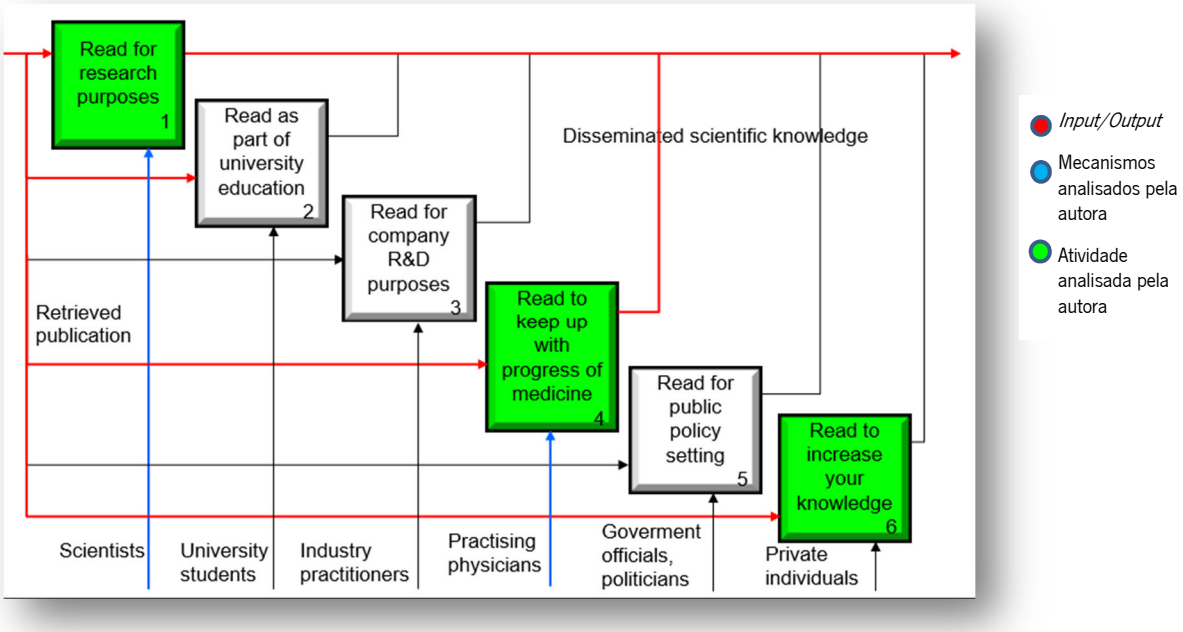


Figura 12. Leitura da publicação (A32342). Recuperado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 38.

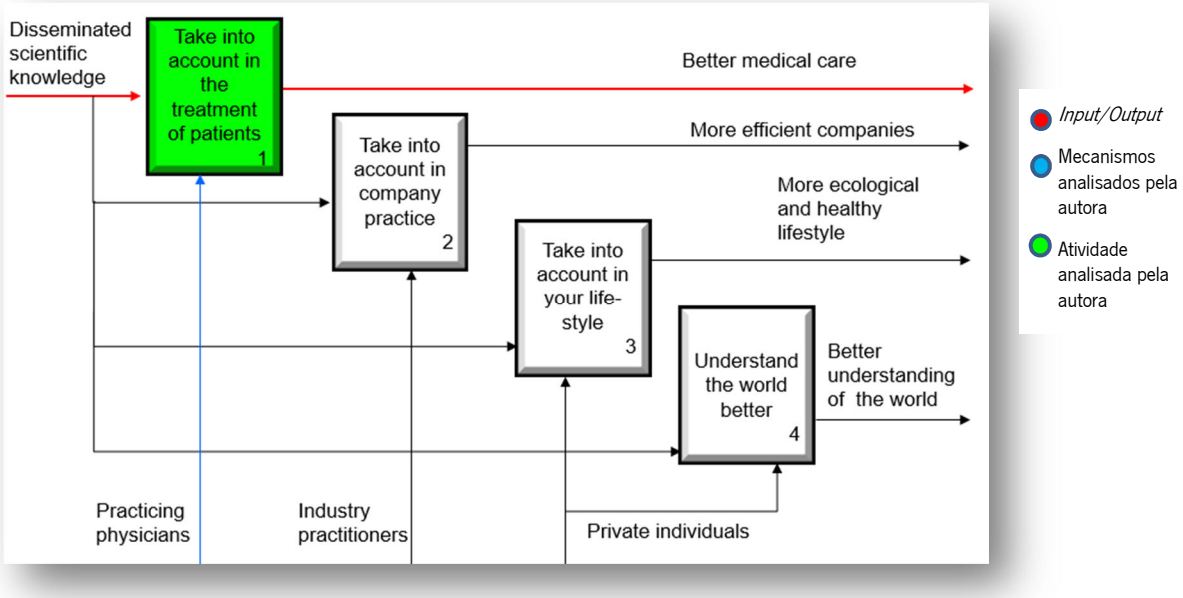


Figura 13. Aplicação na prática (A4). Adaptado de A model of scientific communication as a global distributed information system de B. C. Björk, 2007, *Information Research*, 12, p. 40.

SÍNTESE DA SUBSECÇÃO 2.2

Nesta subsecção foram abordados os tópicos artigos científicos, base de dados bibliográficas e repositórios digitais que contribuem para o processo de comunicação científica que tem sido alterado principalmente a partir do surgimento da Internet e, mais recentemente, pela introdução das tecnologias semânticas no ambiente *Web.*

Björk (2007) apresenta um modelo que traz a comunicação científica como um sistema de informação global. O processo de comunicação científica compreende o conhecimento prévio que, adicionado às motivações científicas e/ou econômicas, induz às novas investigações. Os resultados dessas investigações, ao serem disseminados e aceitos pela comunidade científica, geram novos conhecimentos que podem promover a melhoria da qualidade de vida

Observa-se que o modelo proposto por Björk revela que a comunicação científica:

a) engloba atividades de investigação, publicação e aplicação dos resultados;
b) pode ser viabilizada, no ambiente *Web*, pelas bases de dados e RDCs ou disponibilizadas individualmente na Internet, em acesso livre ou em acesso limitado (publicações pagas);
c) inclui diferentes atores, entre eles, os investigadores, os editores, os financiadores e os utilizadores que exercem papel fundamental no processo; e
d) utiliza tecnologias que potencializam a disseminação do conhecimento científico.

A autora deste estudo se utilizou do modelo de Björk para avançar sobre a comunicação científica, os RDCs e as bases de dados bibliográficas, no sentido de relacionar os RDCs ao *Linked Data* e para sugerir a utilização de TNIs na representação dos recursos por meio de elementos de metadados.

Ao avançar sobre essas questões, sugere-se a adição de um diagrama para o modelo de Björk com o intuito de demonstrar como essas tecnologias atuam no processo de comunicação científica e como a comunicação científica pode se beneficiar com a interoperabilidade semântica e os princípios *Linked Data*.

A próxima subsecção explora o tema necessidades informacionais, que há muito tempo é explorado por vários domínios, sobretudo pela Ciência da Informação, mas que ainda é revestido de grandes desafios. No contexto deste estudo, as NIs são significativas para a qualidade dos RDCs.

2.3 NECESSIDADES DE INFORMAÇÃO

> Cada documento ou registro contém informações sobre algo, e esse algo pode ser determinado objetivamente. De fato, é a especificação do conteúdo que torna possível representar, organizar e armazenar a informação (Choo, 2003, p. 68).

Web Semântica, *Linked Data*, Linguagens de Marcação, *Data Cloud*, publicações enriquecidas semanticamente e repositórios digitais científicos são termos que, se não fazem parte de forma direta do cotidiano dos investigadores e utilizadores no contexto acadêmico/científico, estão por trás das suas atividades de produção científica.

Os profissionais responsáveis por desenvolver tecnologias que têm como propósito capturar, armazenar, organizar e disseminar os dados e as informações, diante de um aumento descontrolado de material disponibilizado na *Web*, estão a buscar soluções em conceitos basilares da Ciência da Informação e Biblioteconomia, como Desenvolvimento de Coleções, Representação Descritiva e Estudos de Usuários.

No entanto, esses profissionais se deparam com dificuldades específicas de um paradigma em que as atividades inerentes à Organização e Representação do Conhecimento não estão mais restritas a um ambiente controlado, como bibliotecas, arquivos e centros de documentação. Os dados, as informações e o conhecimento estão sendo disponibilizados em um ambiente onde as pessoas podem depositar e pesquisar o que desejam, onde o controle dessas atividades é cada dia mais difícil diante da heterogeneidade de sistemas e volume de informação.

No âmbito do desenvolvimento de sistemas de informação e tecnologias ou recursos informacionais, devem ser considerados aspectos relativos ao uso da informação, à necessidade informacional e ao comportamento de busca de informação dos utilizadores que procuram respostas para suas questões, necessidades e motivações (Forsythe, 1998).

O termo necessidade de informação é adotado neste estudo como o "[...] estado ou processo iniciado quando alguém percebe que existe uma lacuna entre a informação e o conhecimento disponíveis para resolver um problema e a sua solução" (Miranda & Tarapanoff, 2008, p. 1, tradução nossa). No entendimento de Prakasan (2013), a necessidade de informação inclui demanda de informação (ou requisitos) e informação requerida (ou desejada) que, embora possam ter diferentes significados, devem ser contempladas por um bom sistema de informação (Smith, 1996). Nessa definição, o termo necessidade de informação refere-se a uma necessidade explicita (Taylor, 1962). Ou seja, quando alguém percebe que existe uma lacuna entre informação e o conhecimento disponível para resolver um determinado problema e a sua solução específica, utilizamos o termo de necessidade de informação.

Neste estudo, o termo "tipo de necessidade de informação" refere-se não a uma necessidade específica, mas a uma abstração sobre um conjunto de necessidades semelhantes sob algum ponto de vista. Assim, quando se refere ao conhecimento médico, os tipos de necessidades de informação podem estar associados a, por exemplo, categorias como tratamento, diagnóstico, entre outras (Allen et al., 2003, Case, 2007, Collins et al., 2009, e Davies, 2008).

Entende-se ainda que as necessidades de informação podem ser satisfeitas a partir dos resultados da busca de informação. Esta representa o esforço em obter a informação em resposta às suas necessidades (Case, 2007). Os tipos de necessidades de informação podem ser referidos como requisitos informacionais quando explicitados para o desenvolvimento de sistemas de informação (Fusco, 2010, Kim, 2010a, Kim & Gilbertson, 2007).

No entanto, mesmo em um âmbito mais restrito, como a produção e a divulgação da ciência, especificamente em relação à comunicação científica, há ausência de respostas objetivas referentes às questões: qual o interesse de um investigador ao buscar um artigo científico ou um documento? Que tipo de informação ele deseja encontrar em uma base de dados bibliográfica de forma a obter documentos mais relevantes para o seu estudo? Como os sistemas estão organizados para atender aos utilizadores com base nos dados disponibilizados para recuperação? Essas questões podem ser desafiadoras para vários domínios do conhecimento.

A necessidade informacional, de acordo com Amato & Straccia (1999) e Courtright (2007), pode variar, dependendo do tipo de utilizador e do contexto em que é analisada ou requerida. Assim, uma das preocupações no desenvolvimento desses estudos é estabelecer o tipo de utilizador a ser analisado, ou seja, sugere-se que o perfil do utilizador seja analisado. Neste estudo, entende-se como perfil do utilizador a "[...] representação estruturada das necessidades dos utilizadores em que um sistema de recuperação deveria se basear para responder às questões desses utilizadores" (Amato & Straccia, 1999, p. 2, tradução nossa).

Forsythe (1998, p. 403, tradução nossa) afirma que é essencial que se faça, antes de desenvolver um sistema de informação, a seguinte pergunta: "O que os utilizadores finais querem em um recurso de informação". Para o autor, a tecnologia não pode resolver sozinha problemas relacionados com a necessidade de informação dos utilizadores; é preciso que os desenvolvedores identifiquem as necessidades e as práticas de suas atividades, para assim fornecer recursos que os auxiliem nessa tarefa.

Para Calva-González (2003), as necessidades têm relação direta com o ambiente de trabalho, com as linhas de investigação em que atuam e com o tópico a ser desenvolvido pelos utilizadores de informação.

A literatura que procura teorizar a NI é vasta (Choo, 2003), assim como a literatura voltada para o comportamento informacional e o uso de informação que por vezes divergem em relação às perspectivas, teorias e modelos. Em alguns estudos, esses tópicos se fundem e, em outros, eles se distanciam. Este estudo, entretanto, não busca mapear todos os estudos passados, mas estabelecer um ponto de referência que possa nortear o desenvolvimento da presente investigação.

Após a análise de estudos sobre as NIs, que decorrem desde a década de 60, com trabalhos que são referências (Menzel, 1964, Paisley, 1968, Taylor, 1962) até os dias atuais, optou-se pela abordagem de um dos precursores nesse tópico, Taylor (1962), que, ao publicar o artigo *The process of asking question*, trouxe novos enfoques em relação à NI que levam em consideração o aspecto emocional e cognitivo e o comportamento na busca e recuperação da informação. O autor apresenta quatro níveis de necessidades: visceral, consciente, formalizada e comprometida.

a) primeiro – a necessidade visceral é real, mas ainda não formulada, por não haver uma consciência clara em relação à necessidade de informação;
b) segundo – corresponde ao estágio em que o utilizador tem a sua necessidade formulada mentalmente (necessidade consciente), mas ainda mal definida. A partir de interações com outras pessoas, consegue torná-la mais clara;
c) terceiro – com a necessidade de informação identificada (necessidade formalizada), consegue explicitá-la, com a expectativa de que o serviço de informação responda prontamente; e
d) quarto – a necessidade comprometida está associada à formulação de questões adequadas à organização do sistema de informação.

A partir dessas abordagens, o autor associa alguns aspectos que interferem na relação homem-máquina: organização do sistema, em que se incluem características de *input* e *output*; tipos, complexidades e características do assunto relacionado com a questão formulada pelo investigador; e a competência do investigador. A organização do sistema está associada à "organização interna" que, no entendimento do autor, corresponde à indexação e aos sistemas de organização do conhecimento que, a seu ver, referem-se ao grau de sofisticação da utilização dos termos, à profundidade de análise, à indexação e ao nível de especificidade.

Nesse sentido, considera-se que, para efeito deste estudo, o investigador tem consciência do que ele deseja, mas depende da "organização interna" dos serviços de informação para que a sua NI seja atendida.

Embora tenhamos encontrado no estudo de Taylor apoio teórico para o desenvolvimento desta investigação, foram evocados outros trabalhos com o intuito de contextualizar a investigação das NIs em algumas situações, visto que, apesar das dificuldades em identificar as necessidades dos utilizadores (Courtright, 2007, Crowston & Kwasnik, 2003, Wilson, 1994, 2000), sabe-se que esse é um ponto crucial para que os serviços de informação desempenhem a função de atender às expectativas dos seus utilizadores de forma eficaz e eficiente.

Borgman (1986) salienta que o processo que envolve a recuperação da informação é complexo por englobar a identificação das NIs dos utilizadores. A NI nem sempre é identificada ou definida; muitas vezes é ambígua, seja em relação à semântica, seja ao sistema utilizado, manual ou informatizado. No sistema informatizado, para que a recuperação da informação possa ser bem-sucedida, são exigidas do utilizador habilidades que envolvem aspectos mecânicos – sintaxe e semântica dos termos utilizados, e conceituais da busca – o "como e o porquê" –, correspondentes ao uso dos pontos de acesso e estratégia de pesquisa.

Alguns autores acreditam que a necessidade informacional pode ser suprida pelo uso dos componentes dos artigos, que representam segmentos lógicos ou físicos de um documento (Bishop, 1998, 1999; Sandusky & Tenopir, 2008).

Nessa perspectiva, Bishop (1998) analisou o sistema DeLIver, da University of Illinois, que permite a pesquisa por termos, em particular por componentes dos documentos, a fim de obter uma informação mais precisa. O *design* do projeto foi baseado na identificação das necessidades dos investigadores para o desenvolvimento de trabalhos e práticas de comunicação, ao utilizarem artigos e infraestrutura digital.

Para Bishop (1998), a indexação de artigos é altamente padronizada, incluindo a identificação de autoria, título, resumo e descritores. Entretanto, a autora identifica que itens existentes nos artigos, mas não explorados nos campos de buscas, podem potencializar a recuperação de documentos mais relevantes. Hjørland, Nielsen, & Williams (2001) também defendem que a combinação do uso das estruturas de documento com as estruturas de informação (*info-brick*) pode levar a melhores resultados de buscas.

A investigação de Bishop (1998, 1999), como a própria autora afirma, está alinhada ao princípio defendido por Paul Otlet, em 1920, que apoia a desagregação dos componentes para que possam ser utilizados individualmente. Otlet propôs que a indexação cobrisse os materiais e partes dos documentos, independentemente do nível de especificidade, de forma a explorar o conteúdo de acordo com o gênero do documento (livro, artigo de periódicos, conferências).

Hjørland, Nielsen e Williams (2001) acrescentam ainda que as diferentes estruturas apresentadas nos distintos tipos de documentos têm consequências na busca. Nos sistemas de recuperação de artigos científicos, a estratégia de busca pode variar com base em questões metodológicas ou conclusões, considerados tópicos de maior interesse. Os autores são favoráveis a que novos estudos sejam realizados com ênfase na objetividade informacional de diferentes pontos de acesso do assunto ou da representação dos documentos (*Subjects Access Points* – SAPs), o que é fundamental para a recuperação deles. Do contrário, se os dados de acesso não forem fornecidos pelos sistemas, os utilizadores não poderão utilizá-los.

Assim como Hjørland, Nielsen e Williams (2001), Taylor (1962) já alertava sobre a necessidade de prover os sistemas com aplicações que permitissem aos utilizadores executar buscas de modo a obter resultados com maior relevância.

Os pontos de acesso ou elementos de metadados são resultados da atividade de indexação de documentos. Vale esclarecer que elementos de metadados, no contexto digital, para alguns autores, possuem o mesmo conceito dos pontos de acesso, ou seja, os pontos de acesso determinam possibilidades objetivas de recuperação de documentos pelos utilizadores, por meio de procedimentos algorítmicos ou automáticos (Hjørland, Nielsen & Williams, 2001). Tais elementos de metadados têm diferentes funções e valores informativos para distintos tipos de pesquisas. A adição de determinados elementos pode beneficiar a recuperação no que diz respeito à precisão.

Markey (2007a, 2007b) considera que especialistas em um determinado domínio buscam alto grau de precisão, delimitando o campo de busca (baseados em elementos de metadados). Suas estratégias buscam identificar pistas constantes em alguma palavra ou frase do **título**, nome de um **autor**, **variável**, um **teste**, um determinado centro de investigação, diminuindo o número de itens recuperados, mas com um alto grau de relevância.

Para Sandusky & Tenopir (2008), os investigadores necessitam de novas formas de localizar, eficientemente, os artigos científicos, assim como meios que lhes possibilitem identificar as partes mais relevantes. Sandusky & Tenopir (2008) associam a dificuldade de muitos investigadores em identificar artigos relevantes à forma como estes são indexados e ao não detalhamento de partes do documento. Os autores fazem referência à ProQuest CSA, que desenvolveu um protótipo de sistema que fornece informações detalhadas por meio da indexação de componentes individuais dos artigos. Esse modelo possibilita a realização de pesquisas boleanas, utilizando autoria, título, **estatística**, **termos geográficos** e taxonômicos. As buscas podem ser limitadas por **mapas**, **figuras**, **fotografias**, por

tipo de disponibilização de artigo (acesso aberto ou restrito, por exemplo) ou artigos que fazem referência a **modelos preditivos**[23].

Dogan, Murray, Névéol, & Lu (2009) registraram, durante um mês, dados coletados na PubMed, a fim de caracterizar as NIs e os hábitos de pesquisa dos seus utilizadores com base no histórico de busca e recuperação de informações. Observaram que as buscas mais frequentes cobrem assuntos como **gene**/**proteína** e **doenças**, em que se utilizam, com frequência, **abreviaturas**, nomes de genes e proteínas. Em vista dos resultados obtidos com o estudo, os investigadores passaram a investigar possíveis estratégias para a melhoria desse panorama. Propuseram um método classificado com base na proximidade do termo e palavras-chave no contexto, com destaque para os termos de busca em títulos de artigos.

Alguns estudos demonstram que o aumento da granularidade das informações nas BDBs contribui para o aprimoramento das buscas, ao contar com o aumento e a diversificação dos elementos de metadados (Bates, 1996, Bishop, 1998, Borgman, 1986, Hjørland, Nielsen, & Williams, 2001, Johnson, Chen, Eng, Makary & Fishman, 2008; Lee & Downie, 2004; Lee, 2013; Sandusky & Tenopir, 2008). A inclusão de elementos de metadados que não descrevem apenas os registros bibliográficos, como título, autor, assunto, possibilita ao utilizador formas de restringir as buscas, otimizando seu tempo dedicado a essa tarefa e obtendo um resultado mais relevante.

Ainda no sentido de explorar e disponibilizar campos de busca, como autor, título e assunto, investigadores como Bishop (1998, 1999) e Crowston & Kwasnik (2003) defendem que é preciso considerar o contexto. Crowston & Kwasnik (2003, p. 4, tradução nossa) afirmam que

> [...] diferentes utilizadores podem exigir diferentes soluções para problemas de informação aparentemente semelhantes, [...] surgem necessidades diferentes porque a situação (ou o contexto) de um utilizador determina não apenas por que os tópicos são solicitados e quais estratégias são invocadas na busca e avaliação de saída, mas também quais tipos de recursos são considerados relevantes e úteis.

Em suma, os investigadores são utilizadores, cuja experiência no desenvolvimento de investigações pressupõe informações com maior nível de especificidade e relevância. Estudos demonstram que, para esse tipo de utilizador, a identificação de artigos relevantes em pouco tempo também tem seus reflexos em determinados tipos de comportamento, como o tempo disponibilizado para leitura de artigos (Berghel,

[23] "É uma função matemática que, aplicada a uma massa de dados, consegue identificar padrões ocultos e prever o que poderá ocorrer" (Taurion, 2014)

Berleant, Foy, & McGuire, 1999, Sandusky & Tenopir, 2008, Shotton, 2009, Tenopir et al., 2009b, Tenopir et al., 2009a).

A análise da literatura permitiu identificar que vários autores abordam a dificuldade de realizar estudos voltados para a identificação de NIs. Essa dificuldade se deve a vários motivos, tais como: uso de procedimentos metodológicos inadequados (Courtright, 2007, Figueiredo, 1983, Freeman Rohde, 1986, Hjørland, Nielsen & Williams, 2001, Kuruppu & Gruber, 2006, Wilson, 1994, 2000); e não inclusão do contexto, para identificação e avaliação das reais necessidades informacionais (Amato & Straccia, 1999, Courtright, 2007, Crowston & Kwasnik, 2003, Rowlands & Nicholas, 2008). Julien & Duggan (2000) salientam que existe uma dificuldade muito grande em identificar e conceituar um contexto. Os aspectos psicológicos, como cognição, são abordados como um outro item de difícil verificação, porque esse processo não é claro para os participantes de uma investigação, para os sujeitos analisados e nem para os investigadores (Crawford, 1978, Crowston & Kwasnik, 2003, Courtright, 2007).

Pode-se observar que, a partir desta revisão, as investigações que procuram identificar as NIs precisam ter uma visão holística do processo de busca e recuperação, pois isso envolve questões relacionadas com psicologia, etnografia e análise de domínio. Procedimentos metodológicos – como o uso de métodos etnográficos ou observacionais – auxiliam na identificação de algumas respostas, porém nem sempre conseguem abstrair a real necessidade de um determinado utilizador ou grupo de utilizadores. Existem variáveis que podem interferir nessa análise, como a questão da temporalidade, que contribui com alterações na forma de o utilizador manifestar suas necessidades, e dos avanços tecnológicos.

Quanto ao quesito avanços tecnológicos, atualmente há tecnologias capazes de aperfeiçoar o processo de busca e recuperação de forma que toda informação possua um significado semântico que pode ser disponibilizado entre os diversos sistemas de informação com o objetivo de atender às necessidades dos utilizadores. Fica sob a responsabilidade dos profissionais que desenvolvem e gerenciam esses sistemas identificar que tipos de informação seus utilizadores procuram e a adoção de procedimentos, como as diretrizes *Linked Data* e a utilização do PAM, para otimizar a eficiência e eficácia dos sistemas de informação, itens abordados nas Subseções 2.1.1 e 2.1.5.

SÍNTESE DA SUBSECÇÃO 2.3

Esta subsecção introduz o tópico necessidade de informação, sob a luz de alguns autores. Trata-se de um tema que vem sendo discutido há algum tempo por estudiosos de diferentes domínios e sob variadas perspectivas. Identifica-se que não há consenso em relação ao seu conceito ou à sua abordagem de

investigação. Observa-se, no entanto, que a identificação dos TNIs de utilizadores deve contemplar a análise do perfil do utilizador, o domínio e o contexto em que atuam. Não é proposta deste estudo uma discussão aprofundada sobre o tópico, mas a identificação de uma linha condutora para o desenvolvimento da presente investigação, restringindo-se, assim, à apresentação de alguns estudos clássicos e/ou norteadores que indicam tópicos referentes às necessidades de informação.

Identificou-se, no estudo de Taylor (1962), a linha condutora para o desenvolvimento deste trabalho em que os sistemas de informação devem proporcionar mecanismos que propiciem aos utilizadores a localização e a recuperação da informação, desde que eles tenham consciência das suas necessidades informacionais. Nessa direção, este estudo buscou identificar as NIs de médicos e investigadores no domínio da Oncologia a fim de moldar essas necessidades a serem denominadas de tipos de necessidades de informação (TNIs).

A Seção 3 apresenta o enquadramento metodológico adotado por esta investigação.

3 ENQUADRAMENTO METODOLÓGICO

> Vamos tentar um método; se falhar, admitimos e tentamos outro. Mas, acima de tudo, tentaremos alguma coisa.
> Franklin Roosevelt

Esta investigação tem carácter interdisciplinar, ao buscar princípios da Ciência da Informação para sustentar o desenvolvimento do artefacto, que está vinculado ao contexto dos Sistemas de Informação, no domínio da Oncologia. Assim, as abordagens de investigação apresentadas procuram contemplar as demandas dos três domínios que suportam a base de desenvolvimento do estudo. No plano inicial do estudo – fundamentação teórica –, foram abordados os tópicos: Web Semântica (*Linked Data*), interoperabilidade (interoperabilidade semântica), metadados, esquemas de codificação, perfil de aplicação, comunicação científica (artigos científicos, RDC, bases de dados bibliográficas) e necessidades informacionais.

O domínio de Sistemas de Informação e suas aplicações

> [...] *is a multidisciplinary subject and addresses the range of strategic*, *managerial and operational activities involved in the gathering*, *processing*, *storing*, *distributing and use of information*, *and its associated Technologies*, *in society and organizations* (UK Academy of Information Systems apud Myers & Avison, 1997, p. 187).

Por outro lado, a Ciência da Informação, é um

> [...] *field devoted to scientific inquiry and professional practice addressing the problems of effective communication of knowledge and knowledge records among humans in the context of social, institutional and/or individual uses of and needs for information. In addressing these problems of particular interest is taking as much advantage as possible of the modern information technology* (Saracevic, 1995, p. 2).

Nesse sentido, estabelece-se a integração Sistemas de Informação e Ciência da Informação com o objetivo de identificar os tipos de necessidades de informação de médicos e investigadores no domínio da Oncologia e o desenvolvimento do perfil de aplicação de metadados para artigos científicos armazenados em repositórios digitais científicos.

Como proposto desde o início do estudo, nesta seção, serão apresentados, primeiramente, conceitos relativos aos tópicos abordados. Dessa forma, procurou-se clarificar a abordagem metodológica adotada para o desenvolvimento desta investigação que envolve dois domínios que nem sempre partilham os mesmos conceitos e terminologias.

Inicia-se pelo termo investigação que, no domínio de Sistemas de Informação, pode ser definido "[...] *as an* ***original investigation*** *undertaken in order to contribute to knowledge and understanding in a particular field*" (Myers, 2013, p. 6, grifo do autor), enquanto método de investigação é considerado "[...] *a strategy of inquiry which moves from the underlying philosophical assumptions to research design and data collection*" (Myers & Avison, 1997, p. 245).

Neste estudo, considerou-se ainda que um processo de investigação "[...] *is the application of scientific method to the complex task of discovering answers (solutions) to questions (problems)* " (Blalock & Blalock, 1982, apud Nunamaker Jr & Chen, 1990, p. 97).

Este estudo insere-se na Design Science, ou seja, nas ciências do artificial, definidas por Simon (1996) como as ciências que se ocupam da conceção de artefactos para um determinado objetivo. Design Science tem por base o desenvolvimento de conhecimento capaz de conceber e desenvolver artefactos (Aken, 2004). Trata-se de uma "ciência projeto - Design Science" que vai além da produção do conhecimento; ela procura aplicá-lo às situações a fim de criar artefactos eficazes (Pries-Heje & Pries-Heje, 2015).

A Design Science possui ainda algumas características que a distinguem de outras metodologias, como o fato de tratar-se de uma investigação baseada na prescrição, que está voltada para a solução de problemas de uma determinada classe de problemas (March & Smith, 1995), ou seja, o Design Science procura criar coisas que possam atender às necessidades humanas (Simon, 1996). A investigação em Design Science é denominada de Design Science Research (DSR) que, de acordo com Simon (1996), deve proporcionar o desenvolvimento do artefacto e permitir a sua avaliação, por meio do rigor e verificação. Isso deve ser viabilizado pela divulgação do resultado para o público adequado.

Conforme o Design Research in Information Systems Group (2006, apud De Villiers, 2012), o DSR não pode ser classificado na linha positivista ou interpretativista de investigação; ele se interpõe entre as duas. O DSR possui uma "[...] abordagem pragmática e resolutiva que tolera a ambiguidade" (tradução nossa). Está alinhado à perspectiva exploratória, aos métodos qualitativos e à prescrição, em que se propõe desenvolver um artefacto ou analisar e compreender o seu funcionamento, dentro de um contexto específico, e às necessidades dos investigados.

A perspectiva exploratória do DSR é "[...] adequada para casos em que ainda não apresentem um sistema de teorias e conhecimentos desenvolvidos [...]"– ou, ainda, "[...] quando se tem uma noção muito vaga do problema de investigação [...]" (Kôche, 1996, p. 126, tradução nossa), em que é preciso aprofundar o conhecimento em relação ao tema tratado para que seja possível identificar questões de

investigações; estabelecer objetivos; nortear o desenvolvimento da investigação; propor soluções para determinado problema ou indagação.

A investigação qualitativa deve-se a fatores como: ter opção de melhores métodos para explorar o comportamento humano; possuir característica exploratória, que proporciona uma melhor investigação de fenômenos complexos quando pouco se sabe sobre eles; e por ser holística (Creswell, 2012, Khoo, Rozaklis, & Hall, 2012, Lakatos & Marconi, 2003). Pode ainda utilizar métodos simples ou múltiplos e diferentes instrumentos de coleta e análise de dados (Fidel, 1993); e ter acesso a registros e a documentos, cuja análise pode subsidiar o desenvolvimento ou melhoria de serviços de informação.

Definida a base epistemológica do estudo, foram identificados, no DSR, aspectos que se encaixam como estratégia metodológica para atingir o propósito deste estudo. O DSR possibilita o desenvolvimento, investigação, validação e avaliação de artefactos como: constructos, modelos, métodos e instâncias de sistema de informações, a fim de resolver novos problemas práticos (Bax, 2014, March & Smith, 1995). Dentre os modelos existentes para o desenvolvimento do DSR (Hevner, March, Park & Ram, 2004, March & Smith, 1995, Offermann, Levina, Schönherr, & Bub, 2009, Peffers, Tuunanen, Rothenberger, & Chatterjee, 2007), optou-se pelo de Hevner (2007), por possibilitar uma visão holística e a interatividade no processo de construção do artefacto a partir da análise cíclica de como este acontece.

Hevner (2007) descreve, por meio de um quadro conceitual, o desenvolvimento de um projeto de investigação no domínio dos Sistemas de Informação, em que procura comprender como ocorre a investigação nesse domínio. O autor define diretrizes para o desenvolvimento de um artefacto que tem como fundamento uma base de conhecimento (fundamentação teórica).

O modelo DSR, concebido por Hevner (2007), possui três fases que contemplam: a percepção, a execução e a avaliação do projeto (Figura 14). As fases são apresentadas em quadros: Ambiente, Design Science Research e Base de Conhecimento. O Ambiente compreende o domínio de aplicação, o problema, os utilizadores envolvidos, o ambiente organizacional e os sistemas tecnológicos. Desse conjunto são extraídos os requisitos e é onde se estabelece o ambiente de teste. O desenvolvimento do artefacto é justificado pelas oportunidades e problemas identificados neste ambiente.

O segundo quadro, Design Science Research, traz os insumos provenientes dos quadros Ambiente e Base de Conhecimento, ou seja, nesta fase são apresentados o artefacto e os métodos para a sua validação, assim como a sua divulgação na forma mais adequada.

O terceiro quadro refere-se à Base de Conhecimento, que proporciona o pilar para o desenvolvimento e validação dos artefactos: fundamentos teóricos, métodos, técnicas e formalismos. A partir da base de conhecimento, estabelece-se o rigor necessário para o desenvolvimento do artefacto e para produção de novos conhecimentos.

Neste modelo, Hevner (2007) demonstra o processo de interação que existe entre as três fases, denominando-o de ciclos. Há uma interação das três fases por meio dos Ciclos da Relevância, do Design e do Rigor. O Ciclo da Relevância, o primeiro ciclo, mostra a interação que existe entre o Ambiente e as atividades de investigação, ou seja, o desenvolvimento do artefacto e dos processos que deve atender às demandas do domínio de aplicação. Para tanto, os resultados devem ser disponibilizados para avaliação pelo seu público-alvo.

O segundo ciclo, do Design, relaciona o desenvolvimento do artefacto e dos processos com a sua validação, que poderá ser feita por especialistas, por meio de publicações avaliadas por pares, testados em laboratório ou submetidos a situações experimentais. O Ciclo do Rigor, como a própria denominação sugere, proporciona rigor à investigação por meio da base teórica e metodológica utilizada para o desenvolvimento do artefacto, base que também será retroalimentada pelo desenvolvimento e divulgação do artefacto.

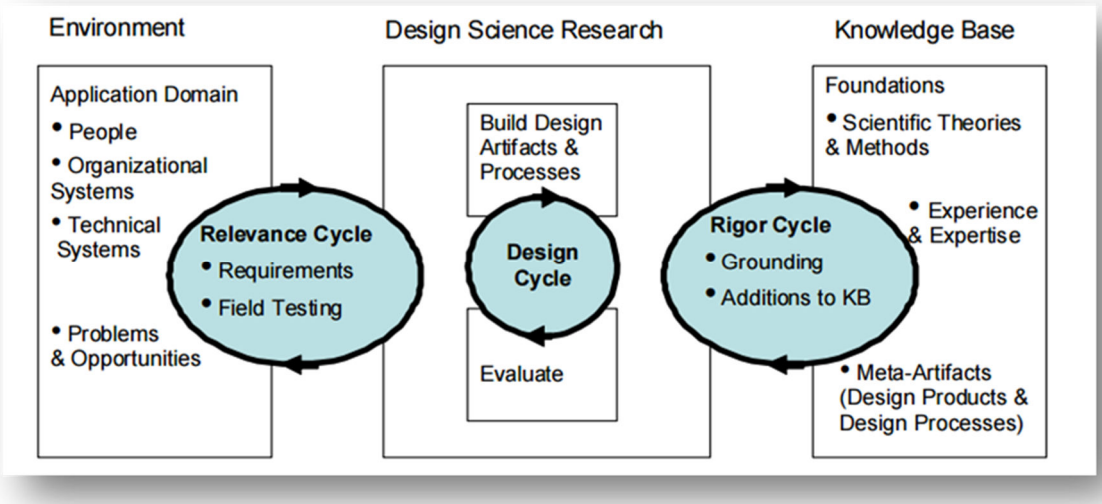


Figura 14. Ciclos do Design Science Research. Recuperado de A three cycle view of design science research de A. R. Hevner, 2007, Scandinavian Journal of Information Systems, 19, p. 4.

Com base nesse modelo, desenvolvemos a estrutura do desenho de investigação deste estudo, como pode ser observado na Figura 15.

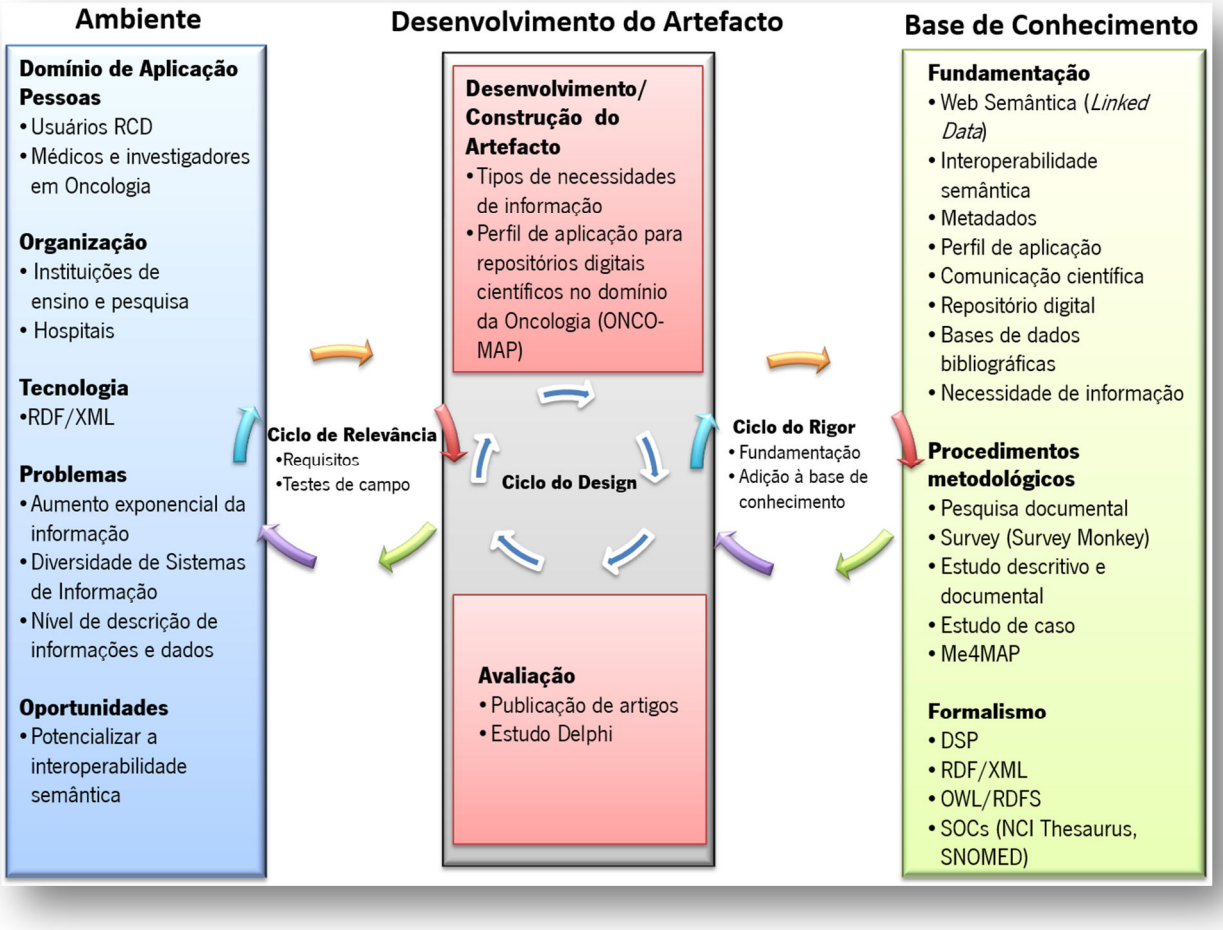


Figura 15. Design Science Research adotado para o desenvolvimento do perfil de aplicação. Adaptado de A three cycle view of design science research de A. R. Hevner, 2007, Scandinavian Journal of Information Systems, 19, p. 4.

Dessa forma, em relação ao Ambiente, o artefacto desenvolvido é direcionado aos atores participantes do contexto dos RDCs no domínio da Oncologia. Trata-se de um domínio complexo que envolve equipes multidisciplinares que produzem e utilizam dados e informações complexas, com aumento exponencial da literatura médica. O contributo advindo desse artefacto tem alcance multidisciplinar e com possibilidade de relevante impacto socioeconômico, já que os RDCs adotam a política de acesso aberto e, com a potencialização da interoperabilidade semântica e alinhamento às diretrizes do LOD, aumentaria substancialmente o número e a qualidade dos recursos a que esses profissionais poderiam ter acesso. Agências de fomento e do governo poderiam aceder à informações que subsidiariam a tomada de decisão em relação ao direcionamento de recursos financeiros, por exemplo, indicadores que poderiam apontar áreas carentes de investigação; por outro lado, os médicos teriam acesso facilitado a recursos informacionais gratuitos de forma eficaz, o que beneficiaria o atendimento de pacientes.

Na fase Ambiente, foram identificados os requisitos[24] para o desenvolvimento do PAM em que se considerou os TNIs de investigadores e médicos que atuam no domínio da Oncologia. Ainda para a identificação dos requisitos, foram adotados métodos e técnicas de coleta de dados fundamentados na análise da revisão da literatura que forneceu base para se optar entre os que melhor se adaptavam ao contexto, ao domínio e ao objetivo do estudo (Base de Conhecimento). A pesquisa documental, o estudo de caso e o *Survey*, com a utilização de questionários eletrônicos, foram selecionados como métodos e técnicas de coleta de dados com base nas exposições apresentadas pelos autores: Creswell (2009, 2012), Jansen (2009), Wildemuth (2016), Myers & Avison (1997) e Myers (2013). Como ferramenta de apoio para organização das fontes selecionadas para o estudo, foi utilizado o Zotero. Para a análise qualitativa dos dados, o NVivo.

Na fase Base de Conhecimento, foi realizado o estudo Estado da Arte sobre os tópicos Web Semântica, interoperabilidade semântica, bases de dados bibliográficas, metadados, perfil de aplicação, comunicação científica, repositórios digitais e necessidade de informação. A análise da literatura apontou lacunas que denotam a necessidade de desenvolvimento do artefacto proposto por esta investigação, bem como a identificação dos TNIs de médicos e investigadores que atuam na Oncologia. Depois da fundamentação teórica, foram elaborados estudos sobre o estágio atual dos PAMs e das bases de dados bibliográficas que contemplam o domínio da Oncologia, cujo resultado proporcionou a identificação de PAMs, esquemas e elementos de metadados utilizados por RDCs; a classificação de pontos de acesso utilizados pelas bases de dados bibliográficas e a ratificação da necessidade de desenvolvimento do artefacto. O desenvolvimento do PAM foi orientado pelo Método Me4MAP, no entanto, sem ter uma preocupação em cumprir todas as etapas e todos os procedimentos de forma rigorosa, visto às particularidades do estudo. Daí um dos aspectos positivos do Me4MAP, de ser um método que trabalha sobre uma perspectiva de um desenvolvimento lógico. Os sistemas de registros de metadados como o Linked open Vocabulary (LOV) e o Metadata Registry foram utilizados para alinhamento dos vocabulários, por se considerar que são tecnologias que possuem fiabilidade, validade e relevância reconhecidas por comunidades de referência na área dos metadados, como a DCMI (http://dublincore.org/) e W3C (http://www.w3.org/). Os formalismos adotados para o desenvolvimento do PAM foram os vocabulários controlados (NCI Thesaurus, SNOMED-CT) e as linguagens RDF/XML, OWL/RDFS, DSP.

No Ciclo do Design, são apresentados os tipos de necessidades de informação de médicos e investigadores e o artefacto – um PAM para artigos científicos no contexto dos RDCs no domínio da Oncologia (Onco-MAP)

[24] Neste estudo considera-se que os requisitos podem estar associados às funcionalidades do sistema (requisitos funcionais, não aprofundados pelo estudo) e às necessidades dos utilizadores (requisitos informacionais).

–, assim como a sua validação. Sendo que a validação do Onco-MAP compreende a confirmação da validade dos elementos de metadados (classes e propriedades definidas a partir dos TNIs) por especialistas da Ciência da Informação que atuam no domínio da Oncologia por meio de um estudo Delphi.

Optou-se pelo método Delphi, que, no estudo desenvolvido por Soares (2010), apresenta-se como uma técnica de comunicação em grupo que busca obter consenso sobre determinado tópico. É interativo, possibilita o anonimato, o *feedback* é controlado e pode ser desenvolvido com o uso de questionários, sem exigir a presença física simultânea dos especialistas.

Embora este estudo compreenda domínios distintos (Tecnologia e Sistemas de Informação e Ciência da Informação), não houve necessidade de incluir especialistas que tenham conhecimento de tópicos como *Linked Data* e interoperabilidade semântica, mas a competência para avaliar o uso de determinados tipos de necessidades de informação na descoberta de recursos. O fato de o Delphi não permitir que sejam avaliadas relações entre os vários tópicos tratados no estudo não inviabiliza o resultado almejado – validar o modelo de dados correspondentes aos TNIs associados aos artigos científicos no domínio da Oncologia (Gordon & Hayward, 1968 apud Soares, 2010). Fizeram parte do Estudo Delphi especialistas que atuam em bibliotecas e instituições no âmbito da Oncologia para validar aspectos relativos ao domínio da Ciência da Informação

SÍNTESE DA SECÇÃO 3

O enquadramento metodológico clarifica conceitos e procedimentos metodológicos empregados para o desenvolvimento da investigação. Descreve-se o Design Science Research (DSR) e a razão pela qual foi adotado para esta investigação, assim como as estratégias e abordagens metodológicas empregadas para a elaboração do artefacto – o PAM. Dentre as etapas desenvolvidas no DSR, serão apresentados nas Seções seguintes os estudos sobre perfis de aplicação e bases de dados bibliográficas (estudos preliminares) e o conclusivo para a elaboração do PAM – a identificação dos TNIs de médicos e investigadores no domínio da Oncologia – e o desenvolvimento do perfil, que correspondem às fases Ambiente e Design, respectivamente.

4 ESTUDOS PRELIMINARES

> *Knowledge is inherently complex and the task of capturing it is correspondingly complex. Thus, we cannot afford to waste whatever knowledge we do succeed in acquiring*
>
> (Neches et al., 1991, p. 54).

Após a revisão de literatura, que contribuiu para o delineamento do objetivo deste estudo – desenvolvimento de um PAM para RDCs o domínio da Oncologia –, foi necessário aprofundar a investigação sobre a elaboração de PAMs no contexto dos RDCs, das publicações científicas e das bases de dados bibliográficas. O estudo procurou identificar a existência de PAMs para domínios específicos e verificar se esses foram desenvolvidos com base nas NIs dos utilizadores.

A representação incompleta ou inadequada do documento e o resultado de buscas que não correspondem às expectativas dos utilizadores são dois problemas apontados por Crowston & Kwasnik (2003) e Björk (2007) relativos à descoberta de recursos em sistemas de informação. O fato de os profissionais da informação realizarem uma representação correta dos documentos nem sempre coincide com o que o utilizador deseja, já que essa representação pode não ser voltada para um contexto ou domínio específico (Hoffman, 2009). Deve-se considerar que o utilizador busca por informações que realmente desejam utilizar.

De acordo com Crowston & Kwasnik (2003), existem fatores que interferem no acesso à informação em relação aos sistemas: as necessidades de informação dos usuários no contexto em que atuam; o conteúdo; e os mecanismos que viabilizam a conexão entre as necessidades e a informação. Segundo os autores, esse mecanismo pode ser uma pessoa, um algoritmo de busca, um ambiente de navegação ou um resumo. Este estudo defende o perfil de aplicação como sendo um mecanismo que pode estabelecer a ponte entre as necessidades e a Informação.

4.1 PERFIL DE APLICAÇÃO DE METADADOS: PANORAMA ATUAL

A utilização dos PAMs vem sendo preconizada já algum tempo pela comunidade de metadados, porém, mesmo assim, nota-se, com base na literatura, que eles não são amplamente utilizados pelos RDCs, seja pelo desconhecimento do que representa e/ou do seus benefícios. Com o objetivo de identificar o que se tem feito em relação aos PAMs, foi desenvolvida a investigação descrita a seguir.

4.1.1 A investigação

Para identificação dos PAMs, foram conduzidas buscas bibliográficas e documentais em bases de dados bibliográficas, a partir de motores de busca, repositórios digitais e também em *sites* de organizações representativas na área, no período de março a abril de 2013, com atualização em agosto de 2016.

Na pesquisa, foram empregados os termos: "*application profile*" e "*scientific publication*"; "*application profile*" e "*digital repositories*"; "*metadata application profile*" ou "*metadata vocabulary*" e "*digital repository*" nas bases de dados Web of Science e Scopus, em que as palavras-chave utilizadas foram introduzidas nos campos de busca "tópico" e "*keywords*", respectivamente. Já em relação às pesquisas realizadas no Google Scholar, adotou-se a opção em "qualquer parte do artigo". Entretanto, no Google Scholar, só foram analisados os 250 resultados mais relevantes, já que, a partir desse número, os títulos/*abstracts* não denotavam relevância em relação ao estudo.

Nas pesquisas realizadas no Google, foram utilizadas as palavras-chave: "*metadata application profile*", "*application profile*", "*metadata schema*", "*metadata vocabulary*" e "*dictionary metadata*" associadas a "*digital repository*", "*bibliographic database*" ou "*digital library*". Essas pesquisas restringiram-se aos idiomas inglês, português e espanhol, sem considerar o período de publicação. Não foi incluído o termo Medicina com a intenção de tornar a pesquisa mais abrangente. Os primeiros resultados já indicavam que a utilização do termo Medicina poderia restringir a identificação de perfis para domínios correlatos.

Além das bases de dados, foram feitas buscas em *sites* como SURF Foundation, Joint Information Science Committee – JISC, Cordis – Seven Framework Program (European Union) e Horizon 2020, Andrew W. Mellon Foundation, National Science Foundation (EUA) e Dublin Core Metadata Initiative.

A seleção dos artigos para análise foi baseada nos títulos e nos *abstracts*. A análise dos artigos buscou identificar informações sobre os perfis de aplicação, que foram classificados como tal, desde que estivessem alinhados com a definição da NISO (2007) e com as recomendações da DCMI (Coyle & Baker, 2009; Nilsson, Baker & Johnston, 2008). A partir da análise, foi possível categorizar os PAMs em relação ao domínio e aos tipos de recursos, com base nas descrições constantes nas publicações científicas e na documentação disponibilizada pelos *sites* das instituições. Utilizou-se o modelo desenvolvido por Manouselis (Manouselis, Najjar, Kastrantas, Salokhe, Stracke, & Duval, 2010) como base para a categorização dos PAMs. O modelo de Manouselis consiste de uma planilha Excel que contém elementos que permitem analisar características dos perfis de aplicação. São ítens

como: informações sobre a identificação do perfil, documentação, escopo, objetivo e esquemas de metadados adotados.

4.1.2 O resultado

O processo de busca possibilitou a identificação de 135 PAMs desenvolvidos para diferentes tipos de recursos e domínios (Tabela 3). Dentre os 135 perfis, foram selecionados para análise do conteúdo apenas os que diziam respeito à comunicação científica (publicações científicas e repositórios digitais científicos), no total de 28. A análise dos perfis selecionados permitiu identificar os esquemas e elementos de metadados utilizados, as propriedades e os esquemas de codificação adotados (Figura 16).

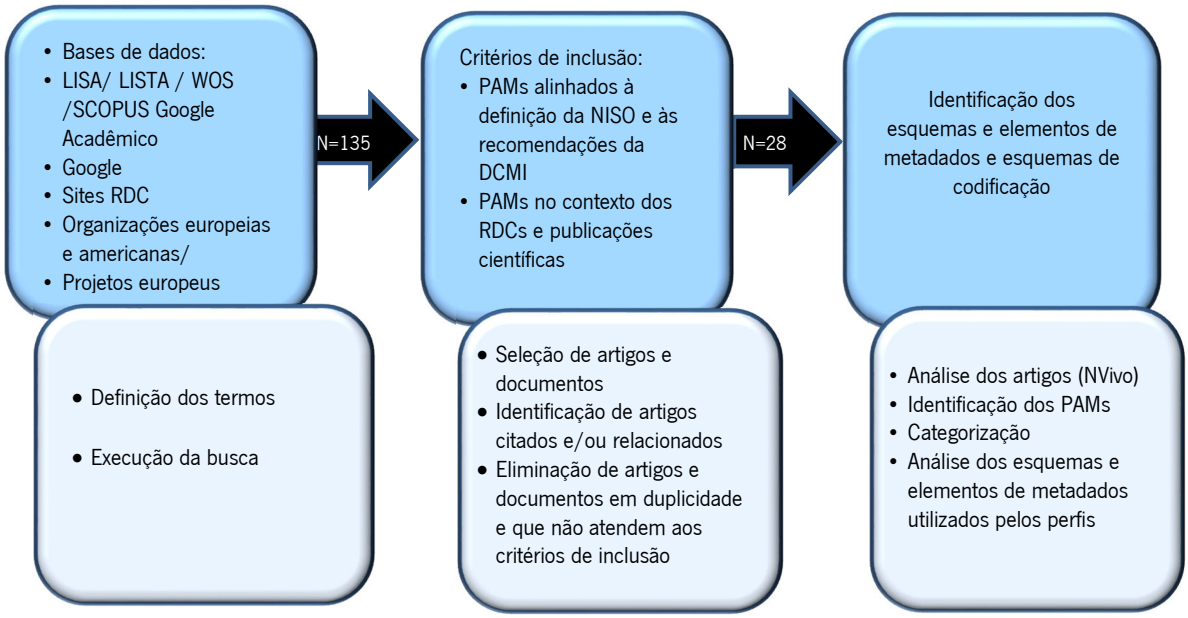


Figura 16. Identificação de Perfis de Aplicação de Metadados.

Tabela 3. **Identificação de perfis de aplicação de metadados**

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Access to Biological Collection Data (ABCD)	Biodiversity Information Standards (TDWG)	2007		Biological Collection Data Banks	Conjunto de dados	Biologia	http://www.tdwg.org/standards/115/
ADN version 0.6.50	Digital Library for Earth System Education		2006	DLESE, GEM, LOM	Objetos educacionais	Meio ambiente	http://www.dlese.org/Metadata/adn-item/0.6.50/docs/framework-documentation.php
Ag-Event AP	Food and Agriculture Organization of the United Nations			DC, Agricultural Metadata Element Set e Australian Government Locator Service Metadata Set	Eventos	Agricultura	http://www.fao.org/3/a-ai255e.pdf
Ag-Job Application Profile	Food and Agriculture Organization of the United Nations			DC TERMS, Agricultural Metadata Element Set,Australian Government Locator Service Metadata Set	Vagas de emprego	Agricultura	http://aims.fao.org/*sites*/default/files/uploadsfiles/ag-job-AP.pdf
Agricultural Learning Resource Application Profile (Ag-LR AP)	AIMS-FAO	2005		DC DC TERMS AgMES	Objetos educacionais	Agricultura	http://aims.fao.org/metadata/sets/ag-lr-application-profile
AGLS Metadata Standard – v. 2.0	National Archives of Australia	2008	2010	DCTERMS AGLS TERMS	Documentos administrativos governamentais	Governo eletrônico	http://www.naa.gov.au/http://www.agls.gov.au/documents/aglsterms/

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
AGRIS-AP	Food and Agriculture Organization of the United Nations	2005		DC, DC TERMS, AGMES	Publicações[25]	Agricultura	http://www.fao.org/docrep/008/ae909e/ae909e00.HTM
ANZ LOM v. 1.13	Educational Services Australia	2002	2015	DC, DCTERMS, SCORM, IEEE LOM	Objetos educacionais	Multidisciplinar	http://www.ndlrn.edu.au/verve/_resources/ANZ_LOM_map.pdf
Application Profile For The Americas Digital Archive	Amanda York Focke, Asst Head, Special Collections Monica Rivero, Digital Project Coordinator	2007		Describing Archives: A Content Standard (DACS) DCMI	Publicações	Multidisciplinar	http://library.rice.edu/about/departments/CDS/cds-projects/Project_profile_OAAP.pdf
Application Profile for Describing CerOrganic DTOs	Agro-Know	2010		IEEE LOM, LRE AP	Objetos educacionais	Agricultura	http://wiki.agroknow.gr/cerorganic/index.php/Main-Page
Asset Description Metadata Schema	Interoperabilty Solutions for European Public Administrations	2012	2013	DCTERMS, FOAF, VCARD Requisitos Funcionais para Registros Bibliográficos -FRBR	Objetos digitais	Governo eletrônico	https://joinup.ec.europa.eu/catalogue/distribution/adms_specification_v100zip
Astronomy visualization metadata (AVM) standard – v. 1.2	Robert Hurt, Lars Lindberg Christensen, Adrienne Gauthier	2008		XMP Standard AVM Standard Exchangeable Image File Format (EXIF) International Virtual Observatory Alliance (IVOA) International Press Telecommunications Council (IPTC)	Objetos multimídia[26]	Astronomia	http://www.virtualastronomy.org/AVM_DRAFTVersion12_rlh02.pdf

[25] Inclui dissertações, teses, anais de congressos, folhetos, livros, monografias, publicações seriadas (revistas científicas) e separatas (Pinheiro & Ferrez, 2014).

[26] Inclui imagens, vídeos e som.

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Badian Collection RUCore Digital Collection	Annamarie Klose	2014		MODS, RULIB	Objetos digitais	Patrimônio cultural	https://rucore.libraries.rutgers.edu/collab/ref/ap_mwg_badian.pdf
Biosci Education Network Metadata Description	Yolanda George, American Association for the Advancement of Science (AAAS)	2003	2011	IEEE LOM	Objetos educacionais	Ciências biológicas	http://www.biosciednet.org/docs/BEN_Metadata_White_Paper.pdf
British Library DOM Legal Deposit eJournal METS Application Profiles	British Library	2007		METS, MODS, PREMIS, JHOVE, British Library's own technical metadata schema for files	Depósito legal	Multidisciplinar	http://www.bl.uk/profiles/ejournal/BL_ejournal_profile.pdf
CanCore learning resource metadata initiative	TeleUniversite and Athabasca University	2000	2006	IEEE LOM	Objetos educacionais	Multidisciplinar	http://cancore.athabascau.ca/en/guidelines.html
CARLCORE Metadata Application Profile	Mark Jordan Lynne McAvoy, Lisa O'Hara, Sharon Ranki, Kathleen Sheare, Nancy Stuart	2008		DC	Publicações	Multidisciplinar	https://www.coffeecode.net/uploads/files/CARLCore_Application_Profile_eng_1-0.pdf
Celebrate Metadata Application Profile	CELEBRATE Project	2003		IEEE LOM	Objetos educacionais	Multidisciplinar	http://celebrate.eun.org/docs/D3_2.pdf
CG LOM Core	Consultative Group on International Agricultural Research – CGIAR	2009		IEEE LOM	Objetos educacionais	Agricultura	https://lirias.kuleuven.be/bitstream/123456789/295784/1/IJMSO%2B41202%2BZschocke%2Bet%2Bal..pdf+&hl=em
CIHC Library Metadata Application Profile	CIHC Library	2008		DC	Publicações	Multidisciplinar	http://www.cihc.ca/files/members/pke/CIHC_Metadata_Template_FINAL_Apr1608.pdf
Comitato Metadati Amministrativi Gestionali – v. 2.0.1	Ministero per I Bene e le Attività Culturali	2001	2009	DC, NISO, X*LINK*	Objetos digitais culturais	Patrimônio cultural	http://www.iccu.sbn.it/opencms/export/*sites*/iccu/documenti/MAG-Reference-201-en.pdf

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
							http://www.iccu.sbn.it/opencms/opencms/documenti/manuale.html#2
Content Standard for Digital Geospatial Metadata, Part 1: Biological Data Profile	USGS Center for Biological Informatics	1999	2010	Necessidades do utilizador FGDC'S Content Standard for Digital Geospatial Metadata.	Conjunto de dados	Biologia	https://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/biometadata/standards/projects/metadata/biometadata/biodatap.pdf
CTDA Connecticut Digital Archive – 3.1	University of Connecticut	2015		MODS	Objetos digitais	Multidisciplinar	http://ctdigitalarchive.org/files/2014/04/CTDA_MAPV3_1.pdf
Darwin core (1.0 and 2.0)	Darwin Core Task Group	2009	2015	DCTERMS	Espécimes biológicas	Diversidade biológica	http://rs.tdwg.org/dwc/terms/index.htm http://rs.tdwg.org/dwc/
Data dictionary – technical metadata for digital still images (ansi/nisoz39.87-2006)	ANSI/ISO	2006		PREMIS DIG35	Recursos multimídia	Património cultural	http://djvu.org/docs/Z39-87-2006.pdf
Datacite Metadata Schema V 3.1 (OpenAIRE)	DataCite - International Data Citation	2015		DCAP DataCite metadata schema	Conjunto de dados	Multidisciplinar	https://schema.datacite.org/meta/kernel-3/index.html
Data*link* Metadata Application Profile – DATMAP v. 3.0	Office of Statistics and Information – Government of Alberta	2014		DCMES	Conjunto de dados	E-Government Estatística	https://osi.alberta.ca/osi-content/Documents/data*link*_Metadata_Application_Profile_FINAL_v3.0_2014-08-21.pdf
DC- Library Application Profile	Rebecca Guenther	2001	2004	DC	Publicações	Multidisciplinar	http://dublincore.org/documents/library-application-profile/
DCAT Application profile for data portals in Europe (DCAT-AP)	Stijn Goedertier	2013		Data Catalogue Vocabulary (DCAT)	Conjunto de dados	Governo eletrônico	https://joinup.ec.europa.eu/asset/dcat_application_profile/description

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
DC-GOV - DCMI Government Application Profile	Maewyn Cumming Palle Aargaard Makx Dekkers	2001		DC, DC-GOV Metadata Element Set, DC-GOV Metadata Element Set Qualifiers	Conjunto de dados	Governo eletrônico	http://dublincore.org/documents/gov-application-profile/
DET Learning Resource Metadata (DETLRM) Application Profile 2.0	State of New South Wales, Department of Education and Training	2010		IEEE LOM	Objetos educacionais	Educação	http://clic.det.nsw.edu.au/clic/documents/detlrm.pdf
Digital Public Library of America - Metadata Application Profile, Version 4.0 (MAP) (DPLA)	Digital Public Library of America - DPLA	2013	2015	Europeana Data Model (EDM) RDF, RDFS OAI-ORE, DCTerms The Basic Geo Vocabulary	Publicações	Multidisciplinar	http://dp.la/info/wp-content/uploads/2015/03/MAPv4.pdf
Dublin Core Application Profile in the Agricultural Domain	DCMI	2001		DC	Objetos digitais	Agricultura	http://aims.fao.org/capacity-development/publications/dublin-core-application-profile-agricultural-domain
DCAP to Analysis Patterns - DC2AP	DCMI	2012		DC, DCTERMS	*Software*	Informática	http://dcpapers.dublincore.org/pubs/article/view/3655/1878
Dublin Core Collections Application Profile	Dublin Core Collection Description Working Group	2007		DC	Publicações	Multidisciplinar	http://dublincore.org/groups/collections/collection-application-profile/
Dublin Core Education Application Profile	DCMI	2007		DC, DCTERMS, IEEE LOM use cases, needs, people, organization	Objetos educacionais	Educação	http://dublincore.org/groups/education/

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
The Dryad Level One Application Profile	DC, Darwin Core DDI, PREMIS	2010		DC DC TERMS Darwin Core	Conjunto de dados	Biology	https://www.nescent.org/wg_digitaldata/Public:Level_One_Application_Profile
eBAnk UK Metadata Application Profile	Traugott Koch	2005		DCMI, DC ABSTRACT MODEL	Objetos digitais	Dados cristalográficos	http://www.ukoln.ac.uk/projects/ebank-uk/schemas/profile/index.html
ECDS profile of ISO 19115:2003 – v. 2.1	Environment climate Data Sweden	2012	2013		Conjunto de dados	Meio ambiente	http://www.smhi.se/polopoly_fs/1.19075!/ECDS%20Metadataprofil%202%201.pdf
Edna Metadata Application Profile 2.0	Australian Information and Communication Technologies in Education Committee (AICTEC)	2007		DCMI, AGLS, IMS Schema	Objetos educacionais	Educação	http://edna.wikispaces.com/edna+metadata+application+profile+2.0
EIAH Metadata Application Profile	Encyclopedia of Iranian Architectural History (EIAH)	2010		DC-LIB, MODS	Objetos digitais	Arquitetura Património cultural	https://www.academia.edu/666306/Lessons_Learned_from_EIAH_Metadata_Application_Profile
Eleonet Metadata Application Profile	European Learning Objects Network	[2007 ?]		IEEE LOM SCORN Metadata AP Requirements	Objetos educacionais	Educação	https://www.medra.org/stdoc/eleonet_metadata_specifications.pdf http://www.eleonet.org/eleonet/area_en-met_schema.htm
EPUB 3 EDUPUB Profile	EPUB Working Group	2015			Objetos educacionais	Educação	http://www.idpf.org/epub/profiles/edu/spec/
eSciDoc Application Profile Publication	Projeto conjunto da Max Planck Society e a FIZ Karlsruhe (Ministério da Educação e Pesquisa Alemão		2015	DC	Publicações	Multidisciplinar	http://colab.mpdl.mpg.de/mediawiki/ESciDoc_Application_Profile_Publication
THE EThOS UKETD_DC Application Profile	British Library			DC, UKETD_DC 'namespace'	Publicações	Multidisciplinar	http://ethostoolkit.cranfield.ac.uk/tiki-index.php?page=The+EThOS+UKETD_DC+application+profile

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Europeana Semantic Elements Specification and Guidelines	Europeana Pro	2013		DC, DCTERMS	Publicações Recursos multimídia	Multidisciplinar	http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/ESE_Documentation/Europeana%20Semantic%20Elements%20Specification%20and%20Guidelines%2014%20July%202013.pdf
EDM – FRBRoo Application Profile	Europeana	2013		DC EDM	Publicações Recursos multimídia	Multidisciplinar	http://pro.europeana.eu/files/Europeana_Professional/Share_your_data/Technical_requirements/ESE_Documentation/Europeana%20Semantic%20Elements%20Specification%20and%20Guidelines%2014%20July%202013.pdf
Facilitating Access to Information on Learning Technology for Engineers Metadata Set	FAILTE	2002		DC, IMS	Objetos educacionais	Engenharia	http://www.failte.ac.uk/documents/guidelines/ index.html
Folklore Description Application Profile		2004		DC, DCTERMS, DC COLLECTIONS, RSLP, MARC, ADL, IEEE LOM	Objetos digitais culturais	Património cultural	
Gateway to educational materials (GEM) element set	ERIC Clearinghouse on Information & Technology Syracuse University	1999		DC Necessidades do utilizador	Objetos educacionais	Educação	http://www7.scu.edu.au/1897/com1897.htm
Grey Highlands Digitization project Metadata Application Profile	Grey Highlands Public Library		2013		Imagem	Multidisciplinar	http://www.greyhighlandspubliclibrary.com/AgnesMacphail/Metadata%20application.htm

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Health Canada *Web* Metadata Application Profile	Health Canada	2008	2009	DC	Publicações	Medicina	http://dev.bowden*Web*.com/html/m/health-canada-meta-data-application-profile-2009.pdf
Images application profile	Joint Information Science Committee (JISC)/ United Kingdom Office for Library and Information Networking (UKOLN/CETIS)		2008	DC IAP MODEL FRBR	Recursos multimídia	Multidisciplinar	http://www.ukoln.ac.uk/repositories/digirep/index/Images_Application_Profile
InterPARES Application Profile	Joseph Tennis University of Washington	2012			Objetos digitais	Metadata	http://www.interpares.org/%5C/ip3/display_file.cfm?doc=ip3_metadata_application_profiles_final_report.pdf
ISDI Metadata Profile – v.1.2	Irish Spatial Data Infrastructure	2013		ISO 19115 ISO 19119	Conjunto de dados Serviços de conjuntos de dados	Geografia	http://www.environ.ie/en/Publications/DevelopmentandHousing/Planning/FileDownLoad,32652,en.pdf
JORUM Application Profile	Adrian Stevenson	2005		UK LOM CORE IEEE LOM	Objetos educacionais	Educaçao	http://www.jorum.ac.uk/wp-content/uploads/2015/01/Jorum_Application_Profile_2005.pdf
KMODDL Application Profile	Cornell University	2004		DCMES, DCMI, MODS, KMODDL	Objetos digitais	http://kmoddl.library.cornell.edu/aboutmeta2.php	http://kmoddl.library.cornell.edu/aboutmeta2.php
Learning Resource Exchange Metadata Application Profile v. 4.7	David Massart and Elena Shulman	2011		IEEE LOM IMS LODE ILOX	Objetos educacionais	Educação	http://lreforschools.eun.org/ *Web*/guest/metadata
LOR Metadata Specification – v.1.2	Nottingham Trent University	2009		DC V.1.1 IEEE LOM V 1.0 GEM 2.0	Objetos educacionais	E-learning	JISCNTU_RA_Metadata_Specification_v1.2 FINAL.doc
LORELOM	SURFNet	2010		IEEE LOM V.1.0	Objetos educacionais	Educação	https://wiki.surfnet.nl/display/standards/LOREIom+Specificatie

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Mathematical Library Application Profile		2010		DCTERMS, DC Collections, PRISM	Publicações	Matemática	http://projecteuclid.org/documents/metadata/mlap/
MEDIN Discovery Metadata Standard	Marine Environmental Data & Information Network	2009	2015	INSPIRE Metadata ISO 19115	Conjunto de dados	Biologia Marinha	http://www.oceannet.org/marine_data_standards/documents/medin_schema_doc_2_3_8.pdf
Metadata Application Profile for FAO'S Learning Resources				IEEE LOM DC	Objetos educacionais	Agricultura	ftp://ftp.fao.org/gi/gil/gilws/aims/metadata/docs/learnap.pdf
Metadata Application Profile for Digital Collections	University of Tennessee at Chattanooga			DCTERMS, DC	Objetos digitais	Patrimônio Cultural	http://wiki.lib.utc.edu/index.php/Metadata_Application_Profile_for_Digital_Collections
METS Profile – Sound Recordings 2 – XML Version of profile document	British Library	2008		METS DOM eJournal ingest DCMI	Recursos multimídia	Multidisciplinar	http://www.bl.uk/profiles/sound/METS_profile.pdf
Michael-eu DCAP	Pete Johnston	2006		DC,DCTERMS, MARC, RSLP, MICHAEL	Publicações	Patrimônio histórico	http://wiki.surf.nl/display/standards/MPEG21+DIDL+Application+Profile+for+Institutional+Repositories
MIMETA Application Profile		2006		IEEE LOM, DC AICC, EDNA, GEM MERLOT	Objetos educacionais	Educação	http://dcpapers.dublincore.org/pubs/article/download/857/853.
Mountain West Digital Library Dublin Core Application Profile	Utah State University Digital Commons@USU	2010		DCMES DCMI	Publicações	Multidisciplinar	http://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=1076&context=lib_pubs

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
MPEG21 DIDL Application Profile for Institutional Repositories v. 3.0	Digital Academic Repository (DARE/SURF)	2009		DC, DCTERMS, DIDL,DII	Objetos digitais	Multidisciplinar	http://wiki.surf.nl/display/standards/MPEG21+DIDL+Application+Profile+for+Institutional+Repositories
National Digital Catalogue Metadata Profile – v1.0	National Catalog Committee	2008		ISO 19115 ISO 19106 Necessidades utilizadores	Conjunto de dados Objetos digitais	Geografia	http://datapreservation.usgs.gov/docs/NGGDPPMetadataProfile.pdf
New Zealand Government Locator Service Metadata Element Set v. 1.2	ANZLIC the Spatial Information Council	2002		DC, DCTERMS, AGLS	Conjunto de dados	Ciência da computação	http://www.anzlic.gov.au/resources/anzlic_metadata_profile
NLM Metadata Schema	US National Library of Medicine	2004		DC	Publicações	Medicina	https://www.nlm.nih.gov/digitalrepository/NLM-DigRep-Requirements-rev032007.doc. http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html
North American Profile of ISO19115:2003 - Geographic information - Metadata (NAP – Metadata, version 1.0.1)	INCITS Technical Committee L1, Geographic Information Systems, and the Canadian General Standards Board – Committee on Geomatics (CGSB-CoG)	2007			Dados digitais geoespaciais	Geografia	http://www.fgdc.gov/standards/projects/incits-l1-standards-projects/NAP-Metadata/napMetadataProfileV101.pdf
North Caroline State and Local Government Metadata Profile for	Noth Caroline Geographic Information Coordinating Council	2014		FGDC Content Standard for Digital Geospatial Metadata, ISO 19115	Dados digitais geoespaciais	Geografia	http://www.nconemap.com/Portals/7/documents/metadata_iso/slgmp.pdf

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Geospatial Data and Services							
NSG Geospatial Core Metadata Profile – NGCMP	National System for Geospatial Intelligence (NSG)	2007		DDMS DC ISO 19115	Conjunto de dados	Geografia	www.gwg.nga.mil/documents/NSG_Geo_Core_MD_Profile.doc
OAI-ORE Perfil de Aplicação	Open Archives Initiative Object Reuse and Exchange	2008		DC, DCTERMS, FOAF, ORE	Objetos digitais	Multidisciplinar	http://www.openarchives.org/ore/ http://www.openarchives.org/ore/terms/
Ohio*LINK* Digital Media Center (DMC) Metadata Application Profile	Ohio University	2006		Anglo-American Cataloguing Rules (AACR2) Archives, Personal Papers, And Manuscripts (APPM) DCMES	Média digital	Multidisciplinar	https://silver.ohio*link*.edu/dms/dmccontribution/DMC_AP.pdf
On Line Journal	Baptista, AA Machado, AB	2001		DCTERMS, SCHEMAS VOCABULARY	Publicações	Multidisciplinar	http://repositorium.sdum.uminho.pt/bitstream/1822/11593/1/200106Baptista%26Machado.pdf
Open Data Metadata Application Profile (ODMAP)	Government of Alberta				Conjunto de dados	Governo eletrônico	http://data.alberta.ca/*sites*/default/files/Open%20Data%20Metadata%20Application%20Profile.pdf
OpenCartable	L'ADULLACT	2004		IEEE LOM, DC	Objetos educacionais	Educação	https://adullact.net/projects/opencartable
OpenGIS Application Profile	Open Geospatial Consortium Inc.	2006		ISO19115/ISO19119 AP	Metadados	Geografia	http://www.opengeospatial.org/standards/requests/35
Organic.Edunet Metadata Application Profile	Organic.Edunet	2009		IEEE LOM	Objetos educacionais	Agricultura	http://wiki.organic-lingua.eu/Organic.Edunet_Metadata_Application_Profile

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Organization Metadata AP	Gauri Salokhe, Valeria Pesce, Joost Liesthout FAO	2008		DC, Agricultural Metadata Element Set E Australian Government Locator Service Metadata Set	Objetos educacionais	Agricultura	ftp://ftp.fao.org/gi/gil/gilws/aims/metadata/docs/organizationap.pdf
OSU Knowledge Bank Metadata Application Profile	Ohio State University Libraries Policies and Procedures Technical Services	2004	2011	Ohio*Link* Metadata Application Profile	Publicações	Multidisciplinar	https://library.osu.edu/documents/knowledge-bank/KnowledgeBankMetadataApplicationProfile2011.pdf
Overheid.nl *Web* Metadata Standaard	Ministerie van Binnenlandse Zaken en Koninkrijksrelaties	2010		DC	Objetos digitais	Governo eletrônico	http://standaarden.overheid.nl/owms/4.0/doc/ http://standaarden.overheid.nl/owms/terms/
Plan4all Metadata Profile – Final version	S. Kafka, R. Fiala	2010		ISO 19139 INSPIRE elements	Conjunto de dados	Plano de ordenamento urbano	https://otik.uk.zcu.cz/bitstream/handle/11025/6215/d3-2-2plan4allmetadataprofile-finalversionrevised-101102043932-phpapp02.pdf?sequence=1
Profilo Applicativo PICO	Scuola Normale Superiore di Pisa	2007	2009	DC	Objetos digitais	Patrimônio cultural	http://www.culturaitalia.it/opencms/export/sites/culturaitalia/attachments/documenti/picoap/picoap1.0.xml
Public Service Information Application Profile	ISA Programme	2008		DC, DCTERMS, AGLS, EGMS, GILS, GOVML	Eventos e serviços públicos	Governo eletrônico	https://joinup.ec.europa.eu/asset/cpsv-ap/description
Quinkan Metadata Application Profile	Quinkan Matchbox Project	2004		DC,DCTERMS, AGLS	Objetos digitais	Património cultural	http://aus*Web*.scu.edu.au/aw03/papers/lissonnet2/
RDN/LTSN LOM Application Profile (RLLOMAP)	JISC/UKOLN	2005	2009	IEEE LOM UK LOM CORE	Objetos educacionais	Educação	http://www.ukoln.ac.uk/projects/iemsr/2wp4/iemsr-site/IEMSR-Static/localhost:8080/iemsr-beta/tabsLayout/AppProfLOMtab.jsp%3Fselected=5.html

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Renardus application profile		2002		DC,DCTERMS, RMES, RMESQ	Publicações	Multidisciplinar	http://dcpapers.dublincore.org/pubs/article/view/650/646
Research Data Management Application Profile	Rice University		2016	DC	Conjunto de dados	Multidisciplinar	https://digitalriceprojects.pbworks.com/w/page/89346902/Research%20Data
RIOXX 2.0	JISC			DC RIOXX	Publicações	Funder and project	http://rioxx.net/v2-0-final/
RSLP Collection Description	UKOLN			DC	Publicações	Multidisciplinar	http://www.ukoln.ac.uk/metadata/rslp/schema/
Scholary Works Application Profile – SWAP	DCMI JISC	2006		FRBR, DC, DCTERMS, FOAF	Publicações	Multidisciplinar	http://www.ukoln.ac.uk/repositories/digirep/index/SWAP
scoLOMFR	Ministére de L'Éducation Nationale, de L'Enseignment Supérieur et de la Recherche	2010		IEEE LOM	Objetos educacionais	Educação	http://eduscol.education.fr/cid57098/le-profil-application-scolomfr.html
SeaDataNet metadata profile of ISO 19115	Enrico Boldrini (CNR-Institute of Atmospheric Pollution Research)	2013		ISO 19115	Conjunto de dados	Biologia Marinha	http://www.seadatanet.org/content/download/20284/140589/file/CDI-profile-V10.0.pdf
Shoreline Metadata Profile of the Content Standards for Digital Geospatial Metadata	Marine and Coastal Spatial Data Subcommittee Federal Geographic Data Committee	2001		FGDC Content Standards for Digital Geospatial Standard, Content Standard For Digital Geospatial Metadata (CSDGM)	Conjunto de dados	Oceanografia	https://www.fgdc.gov/standards/projects/FGDC-standards-projects/metadata/shoreline-metadata/sp_endorsed.pdf
State Library of New South Wales Metadata Application Profile	State Library of New South Wales	2007		DC, METS, MODS,MIX, TEXTMD, METSRIGHS, AGLS	Recursos multimídia Conteúdo *Web*, Conteúdos de Arquivos	Multidisciplinar	http://epubs.scu.edu.au/cgi/viewcontent.cgi?article=1003&context=lib_pubs

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Technical Metadata Application Profile	Aaron Coburn, Esme Cowles DuraSpace	2015		PREMIS EBUCORE	Recursos multimídia e Arquivo *Web*	Multidisciplinar	https://wiki.duraspace.org/display/hydra/Technical+Metadata+Application+Profile
Texas Digital Library Application Profile for ETDs	Surratt, Brian E. Little, Alisha Mitchell, Anne M. Thomale, Jason Flannery, Melinda Reagor	2008		ETD MODS	Publicações	Multidisciplinar	https://scholarship.rice.edu/bitstream/handle/1911/8321/TDLETD.pdf?sequence=7&isAllowed=y
The CARARE metadata schema v.1.1	EUROPEANA CARARE Project	2011		MIDAS Heritage, POLIS DTD	Objetos digitais	Patrimônio histórico	file:///C:/Users/USUARIO/Downloads/CARARE%20metadata%20schema%20outline%20v1.1.pdf
The Common Metadata Elements Standard (CMES)	Government of Ontario	2005		DCMES, DCTERMS	Objetos digitais	Governo eletrônico	https://dr6j45jk9xcmk.cloudfront.net/documents/1888/go-its-46-common-metadata-elements-standard-cmes.pdf
The British Library Application Profile for Objects	The British Library	2005		MODS DC	Objetos digitais	Multidisciplinar	http://www.bl.uk/bibliographic/pdfs/blapv1.pdf
THECB Learning Object Repository Metadata Application Profile	Texas Center for Digital Knowledge – University of North Texas	2007	2009	DCMES DC Education AP IEEE LOM GEM	Objetos educacionais	Educação	http://129.115.100.158/txlor/docs/UNT_MetadataApplicationProfile_02_23_2009.pdf
Time-based Media Application Profile	Gayle Calverley, University of Manchester Pete Johnston, Eduserv	2009		DC, DCTERMS, FOAF, FRBR, EPRINT, MARCREL, TBM	Mídias digitais	Multidisciplinar	http://wiki.manchester.ac.uk/tbmap/images/5/5c/TBMAP_Definitionreport_v3_final.pdf
UK LOM Core	Lorna M. Campbell, Centre for Educational Technology Interoperability Standards	2004		IEEE LOM	Objetos educacionais	Educação	zope.cetis.ac.uk/profiles/uklomcore/uklomcore_v0p2_may04.doc
UK AGMAP 2.1 - UK Academic Geospatial Metadata				ISO 19115 Necessidades dos utilizadores	Conjunto de dados	Geografia	http://www.gogeo.ac.uk/files/metadata/agmap-dataset-profile.pdf

IDENTIFICAÇÃO	DESENVOLVEDOR	CRIAÇÃO	REVISÃO	DESCRIÇÃO	TIPOS DE RECURSOS	DOMÍNIO	URL
Application Profile. V.2							
UK e-Government Metadata Standard – Egms – 3.1		2001	2006	DC	Objetos digitais	e-Government	http://www.esd.org.uk/standards/egms/
UK GEMINI – Geo-spatial Metadata Interoperability Inititative v.2.3	AGI Standards Committee		2015	EU INSPIRE Directive UK GEMINI2	Conjunto de dados	Geografia	http://www.agi.org.uk/join-us/agi-groups/standards-committee/uk-gemini
VET Metadata Application Profile v.1		2007	2014	IEEE LOM	Objetos educacionais	E-learning	http://e-standards.flexiblelearning.net. au/documents/2014-vetadata-application-profile-v1.1.docx
VOA3R AgRes AP - The Virtual Open Access Agriculture & Aquaculture Repository - VOA3R Metadata Application Profile	FAO	2011	2012	DC FRBR SWAP Research Methods Ontology FAO LODE Recommendation 1.0	Publicações	Agricultura e aquacultura	http://ieru.org/voa3r/wiki/images/b/bb/20120913-D3.5-PartB.pdf (Draft 2011)
VRA Core Application Profile		2010		DC VRA	Objetos digitais	Artes visuais	http://www.cs.vu.nl/~schreiber/papers/Assem10a.pdf
VTE Metadata Application Profile – VTE	VTE Metadata Expert Group	2004		IEEE LOM	Objetos digitais Objetos educacionais	Vocacional e Educação técnica (Educação)	http://tesi.fabio.*Web*.cs.unibo.it/twiki/pub/Tesi/DocumentiRitenutiUtili/vetadata_final_v1.02.pdf
WMO Core Profile	World Meteorological Organization	2013	2014	ISO 19115:2003	Conjunto de dados	Geografia	http://www.wmo.int/pages/prog/www/WIS/metadata_en.html

O resultado da análise dos 28 PAMs está explicitado nas Tabelas 4 a 31, em que é possível observar os esquemas, os elementos de metadados e os esquemas de codificação utilizados no contexto da comunicação científica e/ou da Medicina

Tabela 4. **AGRIS-AP**

PERFIL DE APLICAÇÃO - AGRIS–AP	
Esquema de metadados	**Dublin Core/AGLS**
Propriedade	**Esquemas de codificação**
dc:title dc:creator dc:publisher	
dc:date	W3CDTF
dc:subject	ASC CABC DDC LCC UDC AGROVAC CABT ASFAT NALT MESH LCSH
dc:description	
dc:language	ISO639-2 (DCTERMS) ISO639-1 (AGS)
dc:identifier	URI ISBN RN JN PN IPC DOI
agls:citation agls:availability dc:source dc:rights	
dc:coverage	POINT, ISO 3166, TGN, Box, Period, W3CDTF
agls:citation	ISSN CODEN
dc:type	DCMIType
dc:format	IMT
dc:relation	URI (DCTERMS) ISBN (AGS) RN (AGS) JN (AGS) PN (AGS) IPC (AGS) DOI (AGS

Tabela 5. **Application Profile for the Americas Digital Archive**

PERFIL DE APLICAÇÃO – Application Profile for the Americas Digital Archive	
Esquema de metadados	**DC**
Propriedade	**Esquemas de codificação**
dc.contributor.qualifier (author, editor, funder, etc)	
dc.coverage.spatial	
dc.date.digital	
dc.date.issued	
dc.date.original	
dc.description	
dc.description.abstract	
dc.digitization.specifications	
dc.format.medium	
dc.identifier.original	
dc.publisher	
dc.pubPlace	
dc.relation.isPartof (series titles, finding aids, etc)	
dc.relation.isreferencedby (research modules, etc)	
dc.relation.isVersionof (edition statements)	
dc.rights	
dc.source.provenance	
dc.subject.LCSH	
dc.subject.local	
dc.title.alternative	
dc.title.subtitle	
dc.type	
dc.subject.local	
dc.title.alternative	
dc.title.subtitle	
dc.source.collection	

Tabela 6. **CARLCore Metadata Application Profile**

PERFIL DE APLICAÇÃO - CARLCore Metadata Application Profile	
Esquema de metadados	**DC**
Propriedade	**Esquemas de codificação**
dc:coverage	
dc:creator	
dc:date	W3C-DTF
dc:description	
dc:format	MIME Media Types Vocabulary http://www.iana.org/assignments/media-types/.
dc:identifier	
dc:language	RFC 3066 (http://www.ietf.org/rfc/rfc3066.txt) ISO639 (http://www.loc.gov/standards/iso639-2/)
dc:publisher	
dc:relation	
dc:rights	
dc:source	
dc: subject	
dc:type	
dc:title	
dcterms:contributor	

Tabela 7. **CIHC Library Metadata Application Profile**

PERFIL DE APLICAÇÃO -CIHC Library Metadata Application Profile	
Esquema de metadados	**DC**
Propriedade	**Esquemas de codificação**
dc:coverage	
dc:creator	
dc:date	YYYY-MM-DD
dc:description	
dc:format	http://www.iana.org/assignments/media-types/
dc:identifier	
dc:language	
dc:publisher	
dc:relation	
dc:rights	
dc:source	
dc:subject	http://www.nlm.nih.gov/mesh/
dc:type	http://dublincore.org/documents/dcmi-type-vocabulary/
dc:title	
dcterms:contributor	
dcterms:citation	
dcterms:abstract	

Tabela 8. **DC-Library Application Profile**

PERFIL DE APLICAÇÃO - DC-Library Application Profile	
Esquema de metadados	**Dublin Core/MODS**
Propriedade	**Esquemas de codificação**
dc:language	ISO639-2 RFC 1766 RFC3066
dc:coverage	
dcterms:spatial	DCMI Point - http://purl.org/dc/terms/Point ISO 3166 - http://purl.org/dc/terms/ISO3166 DCMI Box - http://purl.org/dc/terms/Box TGN - http://purl.org/dc/terms/TGN
dcterms:temporal	DCMI Period - http://purl.org/dc/terms/Period W3C-DTF - http://purl.org/dc/terms/W3CDTF
dc:rights	URI - http://purl.org/dc/terms/URI
dcterms: audience	
mods:edition	
mods:location	http://www.loc.gov/marc/organizations/ ISO/DIS 15511.3
dc:dcmiType	
dcterms:type	
dcterms:format	
dcterms:medium	
dcterms:identifierCitation	
dcterms:language	
dc:title dc:alternative dc:creator	
dc:contributor	Role list (http://www.loc.gov/...)
dc:publisher	Role list (http://www.loc.gov/...)
dc:subject	DDC - http://www.oclc.org/dewey/ LCC - http://lcweb.loc.gov/catdir/cpso/lcco/lcco.html LCSH - http://lcweb.loc.gov/cds/lcsh.html LCC - http://lcweb.loc.gov/catdir/cpso/lcco/lcco.html LCSH - http://lcweb.loc.gov/cds/lcsh.html MESH - http://nlm.nih.gov/mesh/meshhome.html CDU - http://purl.org/dc/terms/UDC
dc:description dcterms:abstract dc:identifier dcterms:tableOfContents	URI - http://purl.org/dc/terms/URI
dc:date dcterms:created dcterms:valid dcterms:issued dcterms:modified dcterms:dateCopyrighted dcterms:dateSubmitted dcterms:dateAccepted mods:dateCaptured dcterms:available	ISO 8601 (without hyphens) - http://purl.org/dc/terms/ISO8601

PERFIL DE APLICAÇÃO - DC-Library Application Profile	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:type	DCMIType - http://purl.org/dc/dcmitype/ MARC list of sources Art & Architecture thesaurus - http://www.loc.gov/marc/source/aat Thesaurus for graphic materials - http://www.loc.gov/marc/source/gmgpc
dc:format dcterms:medium	IMT http://purl.org/dc/terms/IMT http://www.isi.edu/in-notes/iana/assignments/media-types/media-types
dcterms:extent	
dc:source	URI - http://purl.org/dc/terms/URI
dc:relation	SICI (Serial Item and Contribution Identifier)
dcterms:isVersionOf	ISBN (International Standard Book Number)
dcterms:isFormatOf	ISSN (International Standard Serial Number)
dcterms:hasFormat	DOI (Digital Object Identifier)
dcterms:isReplacedBy	URI - http://purl.org/dc/terms/URI
dcterms:replaces	SICI (Serial Item and Contribution Identifier)
dcterms:isPartOf	ISBN (International Standard Book Number)
dcterms:hasPart	ISSN (International Standard Serial Number)
dcterms:requires	DOI (Digital Object Identifier)
dcterms:isReferencedBy	
dcterms:references	

Tabela 9. **Dublin Core Collections Application Profile**

PERFIL DE APLICAÇÃO - Dublin Core Collections Application Profile	
Esquema de metadados	**Dublin Core/MARCREL/CLD/CDTYPE**
Propriedade	**Esquemas de codificação**
cdtype:CatalogueOrIndex	
cld:associatedCollection	
cld:catalogueOrIndex	
cld:itemType	dcterms:DCMIType, Dublin Core Terms http://purl.org/dc/terms/DCMIType
cld:itemFormat	
cld:dateItemsCreated	
cld:isAccessedVia	
cld:isLocatedAt	
dc:creator	
dc:identifier	
dc:language	dcterms:ISO639-2 http://purl.org/dc/terms/ISO639-2
dc:rights	
dc:subject	dcterms:LCSH, LCC, MESH, UDC
dc:title	
dc:type	dcterms:DCMIType,
dcterms:abstract	
dcterms:accrualMethod	cld:AccrualMethod, Collection Description Terms http://purl.org/cld/terms/AccrualMethod
dcterms:accrualPolicy	cld:AccrualPolicy, Collection Description Terms http://purl.org/cld/terms/AccrualPolicy
dcterms:accessRights	
dcterms:accrualPeriodicity	cld:Frequency, Collection Description Terms http://purl.org/cld/terms/Frequency
dcterms:audience	
dcterms:created	
dcterms:extent	
dcterms:isPartOf	
dcterms:isReferencedBy	
dcterms:provenance	
dcterms:hasPart	
dcterms:spatial	
dcterms:temporal	
dcterms:alternative	
marcrel:OWN	
cld:describedCollection	

Tabela 10. **DPLA Metadata Application Profile 4.0**

PERFIL DE APLICAÇÃO – DPLA Metadata Application Profile 4.0	
Esquema de metadados	**DPLA, DC, EDM, GN, ORE, CNT/GEOJSON/SKOS/RDFS/WGS84_POS**
Propriedade	**Esquemas de codificação**
cnt:chars	
dc:relation	
dc:date	
dc:format	dcterms: imt
dc:rights	
dcterms: identifier	
dcterms: isPartOf	
dcterms: publisher	
dcterms: temporal	
dcterms:alterantive	
dcterms:contributor	
dcterms:creator	
dcterms:description	
dcterms:extent	
dcterms:language	Lexvo
dcterms:rightsholder	
dcterms:spatial	
dcterms:subject	
dcterms:title	
dcterms:type	dcterms:DCMIType
dpla:intermediateProvider	
dpla:isReplacedBy	
dpla:originalRecord	
dpla:providedLabel	
dpla:replaces	
edm:aggregatedCHO	
edm:begin	
edm:dataProvider	
edm:end	
edm:hasView	
edm:isShownAt	
edm:Object	
edm:preview	
edm:provider	
edm:rights	
geojson:geometry	
gn:countryCode	
gn:parentFeature	
oa:annotatedAt	
oa:annotatedBy	
oa:hasBody	
oa:hasTarget	
oa:motivatedBy	oa:motivationScheme
rdfs:comment	
skos:closeMatch	

PERFIL DE APLICAÇÃO – DPLA Metadata Application Profile 4.0	
Esquema de metadados	**DPLA, DC, EDM, GN, ORE, CNT/GEOJSON/SKOS/RDFS/WGS84_POS**
Propriedade	**Esquemas de codificação**
skos:exactMatch	
skos:inScheme	
skos:note	
skos:prefLabel	
wgs84_pos:alt	
wgs84_pos:lat	
wgs84_pos:long	
edm:hasType	AAT
dc:Language	ISO 0639-3

Tabela 11. **EDM – FRBRoo Application Profile**

PERFIL DE APLICAÇÃO – EDM – FRBRoo Application Profile	
Esquema de metadados	**DC, EDM**
Propriedade	**Esquemas de codificação**
dcterms:hasPart	
edm:realizes	
dc:subject	
dc:title	
edm:isRepresentationOf	
dc:Language	
edm:InformationResource	
edm:hasType	
edm:TimeSpan	
edm:occurredAt	
edm:Place	
edm:currentLocation	
rdf:literal	
dc:description	
edm:Agent	
skos:Concept	
edm:hasType	
edm:PhysicalThing	
edm:isDerivativeOf	
edm:incorporates	

Tabela 12. **EThOS UKETD_DC application profile**

PERFIL DE APLICAÇÃO – EThOS UKETD_DC application profile	
Esquema de metadados	**DC, UKETD**
Propriedade	**Esquemas de codificação**
dc:subject	MESH, LCC, LCSH
dc:creator	
dc:identifier	
dc:language	ISO6392
dc:rights	
dc:title	
dc:type	
dcterms:abstract	
dcterms:alterantive	
dcterms:issued	
dcterms:isReferencedBy	
uketdterms:advisor	
uketdterms:authoridentifier	
uketdterms:department	
uketdterms:embargodate	
uketdterms:embargotype	
uketdterms:grantnumber	
uketdterms:institution	
uketdterms:qualificationlevel	
uketdterms:qualificationname	
uketdterms:sponsor	

Tabela 13. **MPEG21 DIDL**

PERFIL DE APLICAÇÃO - MPEG21 DIDL	
Esquema de metadados	**MODS** **Dublin core** **MARCXML**
Propriedade	**Esquemas de codificação**
didl:Descriptor	
didl:Item	
didl:Identifier	
didl:statement	
dcterms:accessRights	
dcterms:available	
dc:description	
dcterms:dateSubmitted	
dc:type	
dcterms:dateModified	

Tabela 14. **eSciDoc Application Profile Publication**

PERFIL DE APLICAÇÃO - eSciDoc Application Profile Publication	
Esquema de metadados	**DC/EDM/RDF/SKOS**
Propriedade	**Esquemas de codificação**
dc:identifier	Literal
dc:language	ISO-639
dc:publisher	Literal
dc:subject	DDC MPINP
dc:title	
dcterms:abstract	Literal
dcterms:alternative	Literal
dcterms:created	http://purl.org/dc/terms/W3CDTF
dcterms:dateSubmitted	http://purl.org/dc/terms/W3CDTF
dcterms:dateAccepted	http://purl.org/dc/terms/W3CDTF
dcterms:issued	http://purl.org/dc/terms/W3CDTF
dcterms:modified	http://purl.org/dc/terms/W3CDTF
escidoc:place	
escidoc:edition	
escidoc:published-online	
dcterms:subject	
escidoc:event	
escidoc:source	
escidoc:location	
escidoc:publication_type	
escidoc:total.numbr-of-pages	
escidoc:academic-degree	
escidoc:tableOfContents	
escidoc:creator	
escidoc:publishing-info	
escidoc:review-method	
escidoc:legal-case	

Tabela 15. **Europeana Semantic Elements Specification and Guidelines**

PERFIL DE APLICAÇÃO – Europeana Semantic Elements Specification and Guidelines	
Esquema de metadados	**DC, EDM**
Propriedade	**Esquemas de codificação**
dc: contributor	
dc: coverage	
dc: creator	
dc: date	ISO 8601
dc: description	
dc: format	
dc: identifier	
dc: language	
dc: publisher	
dc: source	
dc: subject	
dc: title	
dc: type	
dc:relation	
dc:rights	
dctems: tableOfContents	
dcterms:alternative	
dcterms:conformsTo	
dcterms:isPartOf	
dcterms: created	
dcterms: extent	
dcterms: hasFormat	
dcterms: hasPart	
dcterms: hasVersion	
dcterms: isFormatOf	
dcterms: isReferencedBy	
dcterms: isReplaceBy	
europeana: isShownAt	
europeana: isShownBy	
dcterms: issued	
dcterms: isVersionOf	
dcterms: medium	
dcterms: provenance	
dcterms: references	
dcterms: replaces	
dcterms: requires	
dcterms: spatial	
dcterms: temporal	
europeana: dataProvider	
europeana: unstored	
europeana: country	
europeana: isShownAt	
europeana: isShownBy	
europeana: language	
europeana: object	

PERFIL DE APLICAÇÃO – Europeana Semantic Elements Specification and Guidelines	
Esquema de metadados	**DC, EDM**
Propriedade	**Esquemas de codificação**
europeana: provider	
europeana: rights	
europeana: type	
europeana: ugc	
europeana: uri	
europeana: userTag	
europeana: year	
dcterms:replacedBy	
europeana:country	
dcterms:isRequiredBy	

Tabela 16. **Health Canada Application Profile**

PERFIL DE APLICAÇÃO – Health Canada Application Profile	
Esquema de metadados	**DC, HC**
Propriedade	**Esquemas de codificação**
dc: subject	http://clf2nsi2.hcsc.gc.ca/homeaccueil/hccvvcsceng.php
dc:contributor	http://www.tbssct.gc.ca/pubs_pol/sipubs/tb_fip/titles offedorg_e.asp http://direct.srv.gc.ca/cgobin/direct500/BE
dc:creator	http://www.tbssct.gc.ca/pubs_pol/sipubs/tb_fip/titles offedorg_e.asp http://direct.srv.gc.ca/cgobin/direct500/BE
dc.title	
dc.description	
hc.keyword	
dc:language	http://www.loc.gov/standards/iso6392/langhome.html
dc:date	http://purl.org/dc/terms/W3CDTF
dc:title	
hc.owner	
hc.review_date	http://www.w3.org/TR/NOTEdatetime.
dc:type	http://www.tbssct.gc.ca/imgi/mwggtm/typtyp/docs/2003/schem/schem_e.asp
hc.meta_date	
hc.audience	http://www.tbssct.gc.ca/imgi/mwggtm/audaud/docs/2003/schemfinal/schemfinaleng.asp
dcterms: issued	http://purl.org/dc/terms/W3CDTF
dcterms: modified	http://purl.org/dc/terms/W3CDTF
hc:title	
hc:description	

Tabela 17. **Mathematical Literature Application Profile 2010**

PERFIL DE APLICAÇÃO – Mathematical Literature Application Profile 2010	
Esquema de metadados	**DC, Prism, CLD**
Propriedade	**Esquemas de codificação**
cld:isAccessedVia	
dcterms: alternative	
dcterms: bibliographic-Citation	
dcterms: identifier	
dcterms: issued	W3CDTF syntax.
dcterms: abstract	
dcterms: accessRights	
dcterms: contributor	
dcterms: creator	
dcterms: format	IMT vocabulary
dcterms: identifier	
dcterms: isPartOf	Literal
dcterms: language	RFC4646.
dcterms: publisher	
dcterms: references	http://www.w3.org/TR/WCAG//
dcterms: rights	
dcterms: subject	
dcterms: title	
dcterms: type	
prism: doi	
prism: elssn	
prism: endingPage	
prism: isbn	
prism: issn	
prism:number	
prism:publication-Name	
prism:startingPage	
prism:url	
prism:volume	

Tabela 18. **Michael-eu DCAP**

PERFIL DE APLICAÇÃO - Michael-eu DCAP	
Esquema de metadados	**DC**
Propriedade	**Esquema de codificação**
dc:subject	
dc:title	
dc:identifier	
dc:description	
dc:language	
dc:type	
dc:creator	
dc:source	
dcterms:audience	
dcterms:extent	
dcterms:isPartOf	
dcterms:temporal	
rslp: accrualStatus	
dcterms: conformsTo	
dcterms: spatial	
michael:Collection	
michael:DigitalCollection	
dcterms: hasPart	
michael: Period	
michael: illustration	
michael: association	
michael:database	
michael:itemFormat	
michael:isLocatedAt	
michael: isResponsibilityOf	
michael: isAccessedVia	
rslp: administrator	
rslp:isLocationOf	
rslp:legalStatus	
michael:digitisationProcess	
michael:startDate	
michael:endDate	
michael:status	
michael:isFundedBy	
michael:funds	
michael:hasMember	
michael:isMemberOf	
michael:isResponsableFor	
michael:creates	
michael:adminStatus	
michael:street	

PERFIL DE APLICAÇÃO - Michael-eu DCAP	
Esquema de metadados	**DC**
Propriedade	**Esquema de codificação**
michael:tel	
michael:fax	
michael:*e-mail*	
michael:homepage	
michael:administers	
michael:contributesTo	
michael:name	
michael:locality	
michael:protocol	
michael:outputFormat	
michael:accessPoint	
michael:providesAccessTo	
michael:accessCondition	
michael:accessType	
michael:accessibility	
michael:pobox	
michael:postcode	
michael:region	
michael:country	
michael:acronymon	
michael:affiliation	
michael:logo	
michael:accessConditionsDescription	
michael:techRequirements	
michael:interfaceDescription	

Tabela 19. **Mountain West Digital Library Dublin Core Application Profile**

PERFIL DE APLICAÇÃO - Mountain West Digital Library Dublin Core Application Profile	
Esquema de metadados	**DC** **Master Archival Files**
Propriedade	**Esquemas de codificação**
dc:date	W3CDTP
dc:description	
dc:format	Internet Media Type (IMT)
dc:identifier	URI Uniform Resource Identifier http://www.ietf.org/rfc/rfc2396.txt
dc:language	ISO 639-3
dc:publisher	
dc:relation	
dc:rights	
dc:source	
dc: subject	Library of Congress Subject Headings (LCSH) http://www.getty.edu/research/conducting_research / vocabularies/aat/ http://www.nlm.nih.gov/mesh/meshhome.html DC
dc:type	DCMI Type Vocabulary Art and Architecture Thesaurus (AAT) Thesaurus for Graphic Materials (TGM)
dc:title	
dcterms:alterantive	
dcterms:abstract	
dcterms:contributor	Library of Congress Name Authority File (LCNAF) MARC Relator Codes
dcterms:creator	Library of Congress Name Authority File (LCNAF)
dcterms:extent	
dcterms:tableOfContents	
dcterms:temporal	LCSH
dcterms: isPartOf	
conversionSpecifications	
masterChecksum	
masterCompression	
masterCreation	
masterExtent	
masterFormat	
masterMedium	
masterOperatingSystem	
masterQuality	
master*Software*	
Transcription	

Tabela 20. **NLM Metadata Schema**

PERFIL DE APLICAÇÃO – NLM Metadata Schema	
Esquema de metadados	**DC**
Propriedade	**Esquemas de codificação**
dc:format	
dc:language	
dc:publisher	
dc:relation	
dc:rights	
dc:contributor	
dc.subject.keyword	
dc:type	
dc:title	
dc:Title.Alternative	
dc:Subject.Keyword	
dc:Subject.MeSH	
NLMDC.Subject.NLMClass	
NLMDC.Subject.Name	
DC.Date.Issued	W3C-DTF
DC.Date.Created	
NLMDC.Date.Created.Original	
DC.Date.Modified	
NLMDC.Date.Modified.Major	
NLMDC.Date.Expiration	
DC.Identifier.URI	
DC.Identifier.URL	
NLMDC.Identifier.BibUI	
NLMDC.Identifier.ISSN	
NLMDC.Identifier.ISBN	
NLM.Permanence.Level	
NLM.Permanence.Guarantor	
NLMDC.Contributor.Role	
NLM.Mandate	
DC.Description.Notes	
DC.Description.TableOfContents	
DC.Description.Edition	
DC.Description.Abstract	
DC.Description.Audience	
DC.Coverage.Spatial	
DC.Coverage.Temporal	
DC.Relation.IsVersionOf	
DC.Relation.IsFormatOf	
DC.Relation.IsReplacedBy	
DC.Relation.Replaces	
DC.Relation.IsRequiredBy	
DC.Relation.Requires	
DC.Relation.IsPartOf	
DC.Relation.HasPart	
DC.Relation.IsReferencedBy	
DC.Relation.References	

PERFIL DE APLICAÇÃO – NLM Metadata Schema	
Esquema de metadados	**DC**
Propriedade	**Esquemas de codificação**
DC.Contact.*E-mail*	
NLM.Contact.Affiliation	
NLM.Contact.PersonalName	
NLM.ChangeHistory	

Tabela 21. **OAI-ORE**

PERFIL DE APLICAÇÃO - OAI-ORE	
Esquema de metadados	**DC/ORE/DCTERMS/FOAF/RDF/RDFS**
Propriedade	**Esquemas de codificação**
ore:Aggregation	
ore:AggregatedResource	
ore:Proxy	
ore:ResourceMap	
ore:Aggregates	
ore:isAggregatedBy	
ore:describes	
ore:isDescribedBy	
ore:lineage	
ore:proxyFor	
ore:proxyIn	
ore:similarTo	
dcterms:creator	urn:uuid scheme [RFC 4122]
dcterms:modified	ISO8601
foaf:mbox	URI format [RFC 2368]
foaf:name	String Literal
rdf:type	
rdfs:seeAlso	

Tabela 22. **On Line Journal**

PERFIL DE APLICAÇÃO – On Line Journal	
Esquema de metadados	**DC, VCard, IO**
Propriedade	**Esquemas de codificação**
dc:creator	
dc:date	W3C-DTF
dc:description	
dc:format	IMT
dc:identifier	URI
dc:language	RFC 1766
dc:publisher	
dc:relation	URI
dc:rights	
dc:source	URI
dc:subject	
dc:type	DCMI Type Vocabulalry
dc:title	
dcterms:contributor	
vcard: FN	
vcard: *E-mail*	
vcard: URL	
vcard: ORG	
(IO) Document Genre	
(IO) Document State	
(IO) Approval	
(IO) Version Number	
(IO) Document Code	

Tabela 23. **OSU Knowledge Bank Metadata Application Profile**

PERFIL DE APLICAÇÃO - OSU Knowledge Bank Metadata Application Profile	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:coverage	TGN (Getty Thesaurus of Geographic Names), Library of Congress Authority File, Library of Congress Subject Headings, Local Term
dc:creator	Library of Congress Authority File
dc:date	ISO 8601
dc:description	
dc:format	
dc:identifier	
dc:language	ISO639-2
dc:publisher	Library of Congress Name Authority File
dc:relation	
dc:rights	
dc:source	
dc:subject	LCSH, Mesh, TGM
dc:title	
dc:type	
dc:contributor	Library of Congress Authority File

Tabela 24. **Renardus application profile**

PERFIL DE APLICAÇÃO - Renardus application profile	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:creator	
dc:description	
dc:identifier	URI
dc:language	ISO 639-2
dc:subject	LCSH, MeSH, DDC, LCC, UDC Renardus-DDC
dc:title	
dc:type	DCMI Type Vocabulary

Tabela 25. **RIOXX Application Profile**

PERFIL DE APLICAÇÃO - RIOXX	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:format	
dc:title	
dc:identifier	
dc:source	
dc:coverage	
dc:subject	
dc:description	
dc:publisher	
dc:language	ISO 639-3 language codes
rioxxterms:project	
rioxxterms:publication_date	
rioxxterms:type	
rioxxterms:version	
rioxxterms:version-q-record	
rioxxterms:apc	
rioxxterms:author	
ali:free-to-read	
ali:license_ref	
dc:relation	

Tabela 26. **RSLP Collection Description**

PERFIL DE APLICAÇÃO - RSLP Collection Description	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:creator	
dc:description	
dc:format	
dc:identifier	
dc:language	
dc:subject	
dc:title	
dc:type	
cld:strength	
cld:accessControl	
cld:accrualStatus	
cld:legalStatus	
cld:custodialHistory	
cld:note	
cld:hasLocation	
cld:objectName	
cld:agentName	
dcq:spatial	
dcq:temporal	
cld:accumulationDateRange	
cld:contentsDateRange	
cld:owner	
dcq:hasPart	
dcq:isPartOf	
cld:hasDescription	
cld:isDescriptionOf	
cld:hasAssociation	
cld:hasPublication	
cld:accessConditions	
cld:isLocationOf	
cld:seeAlso	
cld:administrator	
cld:address	
cld:postcode	
cld:country	
cld:locator	
vcard:fn	
vcard:org	
vcard:role	
vcard:voice	
vcard:fax	
vcard:*e-mail*	
cld:agentHistory	

Tabela 27. **Scholarly Work Application Profile – SWAP**

PERFIL DE APLICAÇÃO – SWAP	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:type	Eprints EntityType Vocabulary Encoding Scheme
dc:title	Literal Value String
dc:subject	LCSH
dcterms:abstract	
dc:identifier	dcterms:URI
dc:creator	URI of the author and/or to*link* to a *related description*
marcrel:funder	URI of the author and/or to*link* to a *related description*
eprint:grantNumber	Value string
loc:ths	URI of the supervisor and/or to*link* to a *related description*
eprint:affiliated Institution	URI of the affiliated institution and/or to*link* to a *related description*
eprint:hasAdaptation	Value URI
eprint:isExpressedAs	URI of an expression of the eprint and/or to*link* to a *related description*
eprint:status	Eprints Status Vocabulary Encoding Scheme
eprint:version	Value string
dc:language	RFC – 3066
eprint:copyrightHolder	URI of the copyright holder and/or to*link* to a *related description*
dcterms:hasVersion	
eprint:hasTranslation	
dcterms:bibliographicCitation	Guidelines for Encoding Bibliographic Citation Information in Dublin Core Metadata
dcterms:references	
loc:fnd	URI of the editor and/or to*link* to a *related description*
eprint:isManifestedAs	
dc:format	MIME types
dc:language	W3CDTF
dc:publisher	URI of the publisher and/or to*link* to a *related description*
eprint:entity typing	http://purl.org/eprint/entityType/ScholarlyWork http://purl.org/eprint/entityType/Expression http://purl.org/eprint/entityType/Manifestation http://purl.org/eprint/entityType/Copy http://purl.org/eprint/entityType/Person http://purl.org/eprint/entityType/Organization
eprint:status	
marcrel:modified	
eprint:copyrightHolder	
eprint:isAvailableAs	
eprint:accessRights	AccessRights Vocabulary Encoding Scheme
dcterms:licence	Value URI (licence)
dcterms:isPartOf	Value String
foaf:name	foaf:family_name and foaf:givenname

PERFIL DE APLICAÇÃO – SWAP	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
foaf:family_name	foaf:family_name and foaf:givenname for describing persons and foaf:name for describing organisations
foaf:givenname	foaf:family_name and foaf:givenname for describing persons and foaf:name for describing organisations
foaf:workplaceHomepage	Value URI
foaf:mailbox	Mailto URI
foaf:homepage	Value URI

Tabela 28. **Texas Digital Library Application Profile**

PERFIL DE APLICAÇÃO – Texas Digital Library Application Profile	
Esquema de metadados	
Propriedade	**Esquemas de codificação**
mods:titleInfo	
mods:title	
mods:name	
mods:namePart	
mods:roleTerm	
mods:typeOfResource	
mods:genre	
mods:originInfo	
mods:dateCreated	
mods:Language	
mods:languageTerm	
mods:physicalDescription	
mods:form Authority	
mods:InternetMediaType	
mods:abstract	
mods:subject	
mods:topic	
mods:identifier	
mods:extension	
etd:degree	
etd:name	
etd:level	
etd:discipline	
etd:grantor	
etd:degree	
mods:recordInfo	
mods:recordContentSource	
mods:recordCreationDate	
mods:recordIdentifier	

Tabela 29. **VOA3R**

PERFIL DE APLICAÇÃO - VOA3R	
Esquema de metadados	**Dublin Core**
Propriedade	**Esquemas de codificação**
dc:Format	Internet Media Types [MIME].
dc:Language	RFC 4646 [RFC4646].
dc:type	Type – DCMITYPE
dc:title	
dc:creator	
dc:contributor	
dc:publisher	
dc:date	W3CDTF
dc:identifier	
dc:subject	
dc:abstract	
dc:rights	

Tabela 30. **OAI-ORE**

PERFIL DE APLICAÇÃO - OAI-ORE	
Esquema de metadados	**MODS** **DC** **MARC**
Propriedade	**Esquemas de codificação**
ore:Aggregation	
ore:AggregatedResource	
ore:Proxy	
ore:ResourceMap	
ore:Aggregates	
ore:isAggregatedBy	
ore:describes	
ore:isDescribedBy	
ore:lineage	
ore:proxyFor	
ore:proxyIn	
ore:similarTo	
dcterms:creator	urn:uuid scheme [RFC 4122]
dcterms:modified	ISO8601
foaf:mbox	URI format [RFC 2368]
foaf:name	String Literal
rdf:type	
rdfs:seeAlso	

Tabela 31. **Research Data Management Application Profile**

PERFIL DE APLICAÇÃO - Research Data Management Application Profile	
Esquema de metadados	Dublin Core
Propriedade	**Esquemas de codificação**
dc.contributor.org	
dc.contributor.funder	
dc.contributor.	
dc.coverage.point	
dc.coverage.spatial	
dc.coverage.temporal	
dc.creator	
dc.date.created	
dc.date.available	
dc.date.issued	
dc.description	
dc.description.provenance	
dc.description.tableofcontents	
dc.description.version	
dc.digitization.specifications	
dc.format.extent	
dc.format.mimetype	
dc.identifier.citation	
dc.identifier.DOI	
dc.identifier.grantID	
dc.identifier.uri	
dc.language.iso	
dc.publisher	
dc.relation.isreferencedby	
dc.relation.	
dc.rights	
dc.rights.uri	
dc.source.provenance	
dc.subject.	
dc.subject.keyword	
dc.title	
dc.type	

Por meio dessas análises, confirmou-se a inexistência de um perfil nos moldes do proposto por esta investigação – aplicado ao contexto dos RDCs no domínio da Oncologia. Dos 28 PAMs analisados, apenas dois estão relacionados com a Medicina, um no contexto dos documentos digitais e outro no de biblioteca digital. São eles:

a) THE NLM METADATA SCHEMA – o perfil de aplicação desenvolvido para a National Library of Medicine dos Estados Unidos (NLM). O NLM metadata schema é baseado no DC Terms (DCMI) e segue “algumas recomendações do DCLibrary Application Profile (DCLib)”. Em adição ao DC Terms

e DCLib, são utilizados elementos adicionais e propriedades identificadas como requisito pela NLM, por exemplo, o *Permanence ratings*, que é descrito em documento à parte.

O perfil de aplicação descreve os elementos de metadados, propriedades e esquemas de codificação utilizados para descrição dos recursos da NLM. Nele estão incluídos os elementos e qualificadores aprovados pela DCMI, elementos DC com definição de qualificadores pela NLM (NLMDC) e novos elementos definidos pela NLM.

O documento apresenta informações minímas sobre utilização de vocabulário controlado, nome de autoridades ou esquemas de codificação. Fornece informações em relação aos elementos de metadados, definição, obrigatoriedade, ocorrência e alguns comentários sobre os elementos.

b) THE HEALTH CANADA WEB METADATA APPLICATION PROFILE – o Health Canada Web Metadata Application Profile (HC Application Profile) disponibiliza orientações para os desenvolvedores e implementadores de elementos de metadados para descrição de documentos no contexto da saúde nos *sites* canadenses.

Este perfil de aplicação adota o Dublin Core Metadata Specifications under Treasury Board Information Management Standard (TBITS 39.1) e a Parte 1 do Government On-Line Metadata Standard.

O documento descreve o HC Web Metadata Element Set que indica os elementos e os refinamentos. As orientações contêm informações sobre combinação de elementos do DC Terms e de elementos desenvolvidos internamente. O documento traz informações adicionais sobre outros guias, como o Health Canada Metadata Guidelines and Best Practices, que descreve valores dos metadados do Health Canada (disponibilizados mediante autorização para acesso). O Health Canada Application Profile é extensível e utiliza o XML, RDF e *namespace* das espeficações W3C para promover a extensibilidade. É obrigatório o uso de IRI para as páginas Web do Health Canadá.

No documento constam anexos com informações sobre restrições, utilização de vocabulário controlado, nome de autoridades e esquemas de codificação.

O resultado da análise dos dois perfis com base na *framework* desenvolvida por Manousilis é apresentado nas Tabelas 32 e 33.

Tabela 32. **Perfis de aplicação no domínio da Medicina**

1 Informação sobre o perfil de aplicação		
Título do PA	NLM Metadata Schema (AP 1)	Health Canada Web Metadata AP (AP 2)
Organização	National Library of Medicine	Canada Government
Esquema base do PA	DC	DC
Idioma (s) do PA	English	English
IRI do PA	http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html	http://dev.bowdenweb.com/html/m/health-canada-meta-data-application-profile-2009.pdf
Cobertura geográfica do PA	United States	Canada
Comunidade	Bibliographic resources	Web documents
Data de publicação do PA	2002	2008
Data desta análise	2015	2015
2 Documentação		
Modelo conceitual	Não	Não
Infraestrutura técnica	XML	RDF/XML
Declaração de conformidade		Não
3 Escopo e proposta		
Documentação	Não	Sim
Casos de uso	Não	Não
4 Seleção dos elementos		
Elementos obrigatórios		Sim
Elementos adicionais	Sim	Sim
5 Modificações permitidas		
Outras modificações		Não
Modificações não permitidas		Sim
Alteração de elementos existentes		Não
Torna elemento obrigatório em opcional		Não

Tabela 33. **Análise de perfis de aplicação no domínio da Medicina: elementos e propriedades dos metadados**

Atributo / **Esquema**		**Elementos de Metadados**			
		NLM Metadata Schema AP1		Health Canada Web Metadata AP2	
Abstract	DC	DC.Description.Abstract	O	—	
Alternative Title	DC	DC.Title.Alternative	O	—	
Audience	DC	DC.Description.Audience	O	dc.audience	O
Change History	LOC	NLM.ChangeHistory	O	—	
Contact Information, *E-mail*	LOC	DC.Contact.*E-mail*	M	—	
Contact Information, Personal Name	LOC	NLM.Contact.Affiliation	O	—	
Contact Information, Section Affiliation	LOC	NLM.Contact.Affiliation	O	—	
Contributor	DC	DC.Contributor	O	dc.contributor	O
Contributor, Role	LOC	NLMDC.Contributor.Role	O	—	
Coverage	DC	—		dc.coverage	O
Creator	DC			dc.creator	M
Date	DC	—		dc.date	M
Date Created	DC	DC.Date.Created	M	dcterms.issued	M
Date Issued	DC	DC.Date.Issued	M	dcterms.issued	M
Date Modified	DC	—		dcterms.modified	M
Date of last update		DC.Date.Modified	C	—	
Date original form created		NLMDC.Date.Created.Original	O	—	
Edition		DC.Description.Edition	O	—	
Expiration date		NLMDC.Date.Expiration	C	—	
Format	DC	DC.Format	O	—	
Geographic Coverage		DC.Coverage.Spatial	O	—	
Identifier , URI		DC.Identifier.URI	O	—	
Identifier , URL		DC.Identifier.URL	M	—	
Identifier supplied by NLM		NlmUniqueID		—	
Identifier, ISBN		NLMDC.Identifier.ISBN	O	—	
Identifier, ISSN		NLMDC.Identifier.ISSN	O	—	
Identifier, NLM Bib UI		NLMDC.Identifier.BibUI	O	—	
Keyword		DC.Subject.Keyword	O	Keywords	O
Language	DC	DC.Language	M	dc.language	M
Major revision date		NLMDC.Date.Modified.Major	O	—	
Mandate		NLM.Mandate	O	—	
Metadata Date		—		meta_date	C
Owner		—		owner	O
Permanence Guarantor		NLM.Permanence.Guarantor	M	—	
Permanence Level		NLM.Permanence.Level	M	—	
Publisher	DC	DC.Publisher	M	—	
Relation	DC	DC.Relation	O	—	
Relation, Has Part	DC	DC.Relation.HasPart	O	—	
Relation, Is Format of		DC.Relation.IsFormatOf	O	—	

Atributo	Esquema	Elementos de Metadados			
		NLM Metadata Schema AP1		Health Canada Web Metadata AP2	
Relation, Is Part Of	DC	DC.Relation.IsPartOf	O	—	
Relation, Is Referenced By	DC	DC.Relation.IsReferencedBy	O	—	
Relation, Is Replaced By	DC	DC.Relation.IsReplacedBy	O	—	
Relation, References	DC	DC.Relation.References	O	—	
Relation, Requires		DC.Relation.Requires	O	—	
Relation, Version of		DC.Relation.IsVersionOf	O	—	
Review Date		—		review_date	C
Rights	DC	DC.Rights	M	—	
Subject	DC	dc.Subject.Keyword	O	dc.subject	M
Subject, Class Number		NLMDC.Subject.NLMClass	O	—	
Subject, MeSH		DC.Subject.MeSH	O	—	
Subject, Name		NLMDC.Subject.Name	O	—	
Subject, Title		NLMDC.Subject.Title	O	—	
Table of Contents	DC	DC.Description.TableOfContents	O	—	
Title	DC	dc.title	M	dc.title	M
Type	DC	DC.Type	M	dc.type	O
		ResourceInfo		dc.type (extended)	O

Nota. Obrigações: M – mandatório, O – Opcional, C – Condicional. Recuperado de http://dev.bowdenweb.com/htmlm/health-canada-meta-data-application-profile-2009.pdf e http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html

A análise dos dois perfis indica que as duas instituições fazem uso de esquemas de codificação, entretanto o NLM Metadata Schema adota apenas o SES para datas (Tabela 34).

Tabela 34. **Análise dos perfis de aplicação no domínio da Medicina: esquemas de codificação**

Atributo	**Esquema de Codificação**	
	NLM Metadata Schema AP1	Health Canada Web Metadata AP2
Abstract	—	—
Alternative Title	—	—
Audience	—	http://www.tbs-sct.gc.ca/im-gi/mwg-gtm/audaud/docs/2003/ /schemfinal-eng.asp
Change History	—	—
Contact Information, *E-mail*	—	—
Contact Information, Personal Name	—	—
Contact Information, Section Affiliation	—	—
Contributor	—	http://www.tbssct.gc.ca/pubs_pol/ sipubs/tb_fip/titlesoffedorg_e.asp
Contributor, Role	—	—
Coverage	—	—
Creator		http://www.tbssct. gc.ca/pubs_pol/sipubs/tb_fip/titles offedorg_e.asp) http://direct.srv.gc.ca/cgo-bin/direct500/BE
Date	—	http://purl.org/dc/terms/W3CDTF
Date Created		http://purl.org/dc/terms/W3CDT
Date Issued	http://purl.org/dc/terms/W3CDTF	—
Date Modified	http://purl.org/dc/terms/W3CDTF	http://purl.org/dc/terms/W3CDTF
Date of last update	—	—
Date of revised	—	—
Date original form created	—	—
Edition	—	—
Expiration date	—	—
Format	—	—

Recuperado de http://dev.bowdenweb.com/htmlm/health-canada-meta-data-application-profile-2009.pdf e http://www.nlm.nih.gov/tsd/cataloging/metafilenew.html.

A análise dos 28 PAMs adotados pelos RDCs e/ou para descrição de publicações científicas possibilitou a categorização dos elementos mais utilizados nos 28 PAMs adotados pelos RDCs e/ou para descrição de publicações científicas (Tabela 35). Dentro de cada categoria, foram sombreadas as propriedades que se apresentaram com maior frequência de uso em relação ao seu objetivo. A Tabela 36 mostra os esquemas de codificação descritos nos PAMs.

Tabela 35. **Classificação dos elementos de metadados registrados nos PAMs**

Categoria	Atributo	Propriedade[27]	Frequência de uso
Título	Title	dc:title	20
		dcterms:title	3
		hc:title	1
	Alternative Title	dc:Title.Alternative	1
		dcterms:alternative	8
		dc.title.subtitle	1
		mods:titleInfo	1
		mods:title	1
Responsabilidade	Creator	dc:creator	17
		dcterms:creator	6
		escidoc:creator	1
		michael:name	1
		rioxxterms:author	1
		cld:owner	1
		foaf:family_name	1
		foaf:firstName	1
		foaf:givenname	1
		foaf:mailbox	1
		foaf:mbox	1
		foaf:name	1
		etd:name	1
		mods:name	1
		mods:namePart	1
		mods:roleTerm	1
		Author	1
	Contributor	dc:contributor	8
		dcterms:contributor	1
		uketdterms:advisor	1
		dc.contributor.qualifier	1
	Publisher	dc:publisher	14
		dc.pubPlace	1
Características físicas	Date	DC.Date.Created	1
		dc.date.original	1
		mods:dateCreated	1
		DC.Date.Issued	2
		DC.Date.Modified	1
		mods:recordCreationDate	1
		dc:date	13
		dc.date.digital	1
		cld:accumulationDateRange	1
		cld:contentsDateRange	1
		dcterms:created	4
		dc.digitization.specifications	1
		dcterms:dateAccepted	2
		dcterms:dateCopyrighted	1
		dcterms:dateModified	1
		dcterms:dateSubmitted	3
		dcterms:issued	5
		dcterms:modified	4
		mods:extension	1

[27] Foi mantida a grafia identificada nas fontes.

Categoria	Atributo	Propriedade[27]	Frequência de uso
		dcterms:valid	1
		edm:begin	1
		edm:dataProvider	1
		edm:end	1
		cld:dateItemsCreated	1
		hc.meta_date	1
		hc.review_date	1
		michael:endDate	1
		michael:Period	1
		michael:startDate	1
		mods:dateCaptured	1
		NLMDC.Date.Created.Original	1
		NLMDC.Date.Expiration	1
		NLMDC.Date.Modified.Major	1
		rioxxterms:publication_date	1
		uketdterms:embargodate	1
	Identifier	DC.Identifier.URI	1
		dc.identifier.original	1
		mods:originInfo	1
		dc:identifier	19
		edm:aggregatedCHO	1
		mods:recordIdentifier	1
		NLMDC.Identifier.BibUI	1
		NLMDC.Identifier.ISBN	1
		NLMDC.Identifier.ISSN	1
		prism:eIssn	1
		mods:identifier	1
		prism:isbn	1
		prism:issn	1
		prism:url	1
		VCard:url	1
		uketdterms:authoridentifier	1
		mods:form authority	1
	Language	dcterms:language	4
		dc:language	21
		mods:language	1
		mods:languageTerm	1
	Format/medium	cld:itemFormat	1
		dc.format.medium	1
		dc:format	16
		dcterms:format	2
		dcterms:hasFormat	2
		dcterms:isFormatOf	1
		michael:accessCondition	1
		michael:accessibility	1
		michael:accessPoint	1
		michael:accessType	1
		michael:DigitalItem	1
		michael:interfaceDescription	1
		michael:outputFormat	1
		edm:TimeSpan	1
		michael:protocol	1
		michael:providesAccessTo	1
		michael:techRequirements	1

Categoria	Atributo	Propriedade[27]	Frequência de uso
	Edition/version	eprint:status	1
		mods:edition	1
		rdf:literal	1
	Size	dcterms:Extent	6
Localização	Availability	dc.Coverage.Spatial	2
		dcterms:available	4
		escidoc:location	1
		eprint:isAvailableAs	1
		michael:InstitutionLocation	1
		michael:Place	1
		escidoc:location	1
		cld:locator	1
		cld:postcode	1
		cld:country	1
		michael:street	1
		edm:occurredAt	1
		edm:Place	1
		edm:currentLocation	1
		michael:pobox	1
		michael:locality	1
		michael:postcode	1
		michael:region	1
		michael:country	1
		mods:location	1
Assunto	Subject	dc:subject	19
		dc.subject.keyword	1
		dcterms:subject	4
		hc.keyword	1
		mods:topic	1
		etd:discipline	1
		cld:agentHistory	1
		cld:agentName	1
		dc.subject.local	1
		skos:Concept	1
		cld:objectName	1
		mods:subject	1
	Coverage	DC.Coverage.Temporal	1
		dc:coverage	9
		dcq:temporal	1
		dcq:spatial	1
		dcterms:spatial	5
		dcterms:temporal	6
Conteúdo	Description	cld:hasDescription	1
		cld:accrualStatus	1
		cld:catalogueOrIndex	1
		cld:custodialHistory	1
		cld:isDescriptionOf	1
		cld:note	1
		etd:degree	1
		etd:level	1
		europeana:userTag	1
		escidoc:total.numbr-of-pages	1
		cld:strength	1
		edm:InformationResource	1

Categoria	Atributo	Propriedade[27]	Frequência de uso
		DC.Description.Abstract	1
		DC.Description.Audience	1
		DC.Description.Edition	1
		DC.Description.Notes	1
		DC.Description.TableOfContents	1
		eprint:isExpressedAs	1
		escidoc:tableOfContents	1
		mods:form authority	1
		dc:abstract	2
		dc:audience	2
		mods:abstract	2
		dc:description	16
		dcterms:abstract	8
		dcterms:audience	4
		dcterms:description	1
		dcterms:tableOfContents	3
		eprint:grantNumber	1
		hc:description	1
		michael:funds	1
		michael:isFundedBy	1
		michael:Programme	0
		michael:Project	0
	Type	cld:itemType	1
		mods:typeOfResource	1
		dc:dcmiType	1
		dc:type	19
		mods:internetMediaType	1
		mods:genre	1
		dcterms:type	3
		edm:hasType	2
		eprint:EntityType	1
		escidoc:publication_type	1
		europeana:type	1
		michael:Collection	1
		michael:database	1
		michael:DigitalCollection	1
		michael:itemFormat	1
		michael:productService	1
		rdf:type	1
		edm:PhysicalThing	1
		mods:physicalDescription	
		rioxxterms:type	1
		edm:hasType	1
Direitos autorais	Rights Description	ali:free-to-read	1
		ali:license_ref	1
		uketdterms:embargotype	1
	Access rights	cld:accessConditions	1
		cld:accessControl	1
		cld:isAccessedVia	1
		cld:legalStatus	1
		dc:rights	14
		dcterms:accessRights	4
		dcterms:licence	1
		dcterms:rights	1

Categoria	Atributo	Propriedade[27]	Frequência de uso
		dcterms:rightsholder	1
		edm:rights	1
		eprint:accessRights	1
		eprint:copyrightHolder	1
		europeana:rights	1
		michael:accessCondition	1
		michael:accessConditionsDescription	1
		michael:accessibility	1
		michael:accessType	1
		rslp:legalStatus	1
Relações	Relation	cld:associatedCollection	1
		cld:describedCollection	1
		cld:hasAssociation	1
		cld:hasLocation	1
		cld:hasPublication	1
		cld:isLocatedAt	1
		cld:isLocationOf	1
		cld:seeAlso	1
		DC.Relation.HasPart	1
		DC.Relation.HasPart	1
		DC.Relation.IsFormatOf	1
		DC.Relation.IsPartOf	1
		DC.Relation.IsReferencedBy	1
		DC.Relation.IsReplacedBy	1
		DC.Relation.IsRequiredBy	1
		DC.Relation.IsVersionOf	1
		DC.Relation.References	1
		DC.Relation.Replaces	1
		DC.Relation.Requires	1
		dc.relation.isreferencedby	1
		edm:isRepresentationOf	1
		dc:relation	12
		edm:incorporates	1
		edm:isDerivativeOf	1
		dcq:hasPart	1
		dcq:isPartOf	1
		dcterms:hasPart	5
		dcterms:hasVersion	2
		dcterms:isPartOf	8
		dcterms:isReferencedBy	4
		dcterms:isReplacedBy	1
		dcterms:isVersionOf	2
		dcterms:references	4
		dcterms:replacedBy	1
		dcterms:replaces	1
		dcterms:requires	0
		dpla:isReplacedBy	0
		dpla:replaces	2
		edm:hasView	1
		eprint:hasTranslation	1
		eprint:version	1
		michael:accessPoint	1
		michael:interfaceDescription	1
		michael:outputFormat	1

Categoria	Atributo	Propriedade[27]	Frequência de uso
		michael:protocol	1
		michael:providesAccessTo	1
		michael:techRequirements	1
		rioxxterms:version	1
		rioxxterms:version-q-record	1
		eprint:hasAdaptation	1
		eprint:isExpressedAs	1
		eprint:isEManifestedAs	1
Citação	Citation	agls:citation	1
		dcterms:bibliographicCitation	3
		dcterms:citation	1
		dcterms:identifierCitation	1
		Eprint:affiliatedInstitution	1
Instituicional	Institution	cld:address	1
		michael:acronym	1
		michael:administers	1
		michael:adminStatus	1
		michael:affiliation	1
		michael:Agent	1
		edm:Agent	1
		michael:contributesTo	1
		dc.contributesTo	1
		michael:creates	1
		michael:email	1
		michael:fax	1
		michael:hasMember	1
		michael:homepage	1
		etd:grantor	1
		michael:isMemberOf	1
		michael:isResponsibleFor	1
		michael:logo	1
		michael:name	1
		michael:tel	1
		uketdterms:department	1
		uketdterms:institution	1
		vcard:org	2
		foaf:workplaceHomepage	1
		foaf:homepage	1
Coleção	Collection	dc:source	11
		eprint:source	1
		mods:recordInfo	1
		mods:recordContentSource	1
		dc.source.collection	1
		dc.source.provenance	1
		dcterms:accrualMethod	2
		dcterms:accrualPeriodicity	2
		dcterms:accrualPolicy	2
		edm:object	1
		edm:isShownAt	1
		edm:preview	1
		edm:provider	1
		cld:administrator	1
		escidoc:review-method	1
		escidoc:publishing-info	1

Tabela 36. **Esquemas de codificação descritos nos PAMs**

Categoria	Regra/Vocabulário	Namespace
Tipo de recurso	DCMITYPE	http://purl.org/dc/dcmitype/
	MARC list of sources	http://www.loc.gov/marc/source/aat
		http://www.loc.gov/marc/source/gmgpc
		info:eu-repo/semantics/
Gênero	Art and Architecture Thesaurus (AAT)	
	Thesaurus for Graphic Materials (TGM)	
Idioma	RFC 4646	http://purl.org/dc/terms/ISO639-2
	ISO 639-2	http://www.loc.gov/standards/iso6392/langhome.html
	RFC 1766	
	ISO639-1	
	RFC – 3066	
Assunto	AGROVAC	
	ASFAT	
	CABC	
	CABT	
	CODEN	
	DDC	DDC - http://www.oclc.org/dewey/
	LCC	LCC - http://lcweb.loc.gov/catdir/cpso/lcco/lcco.html
	LCSH	LCSH - http://purl.org/dc/terms/LCSH
	MESH	MESH - http://nlm.nih.gov/mesh/meshhome.html
	MPINP	
	NALT	
	Renardus-DDC ASC	CDU - http://purl.org/dc/terms/UDC
	UDC	
Data	W3CDTF	http://purl.org/dc/terms/W3CDTF
	ISO8601	http://purl.org/dc/terms/ISO8601
Formato	Internet Media Types [MIME].	http://www.iana.org/assignments/media-types/
	IMT	http://purl.org/dc/terms/IMT
		http://www.isi.edu/in-notes/iana/assignments/media-types/media-types
		Eprints EntityType Vocabulary Encoding Scheme
		http://dublincore.org/documents/dcmi-type-vocabulary/
		http://www.tbssct.gc.ca/imgi/mwggtm/typtyp/docs/2003/schem/schem_e.asp
Localização	ISO/DIS 15511.3	DCMI Point - http://purl.org/dc/terms/Point
		ISO 3166 - http://purl.org/dc/terms/ISO3166
		DCMI Box - http://purl.org/dc/terms/Box
		TGN - http://purl.org/dc/terms/TGN
		Marc - http://www.loc.gov/marc/organizations/
Identificadores	DOI	
	IPC	
	ISBN	
	ISSN	
	PN	
	RN	
	SICI	
	URI	http://purl.org/dc/terms/URI
Direitos autorais		info:eu-repo/semantics/AccessRights Vocabulary Encoding Scheme
Responsabilidade		Library of Congress Authority File
		http://www.tbssct.gc.ca/pubs_pol/sipubs/tb_fip/titlesoffedorg_e.asphttp://direct.srv.gc.ca/cgobin/direct500/BE

Em complementação a esta análise, foi desenvolvido um *Survey*, com aplicação de questionário eletrônico, para os administradores de repositórios digitais, a fim de identificar a existência de outros perfis não localizados na pesquisa bibliográfica e documental (Survey para identificação de perfis de aplicação adotados por repositórios científicos). O resultado do *Survey* é apresentado na subsecção seguinte.

4.2 *SURVEY* PARA IDENTIFICAÇÃO DE PERFIS DE APLICAÇÃO ADOTADOS POR REPOSITÓRIOS CIENTÍFICOS

Em complementação ao estudo anterior, com o objetivo de aprofundar a identificação da utilização de PAMs por parte dos RDCs, adotou-se o *Survey* com a utilização do questionário eletrônico *SurveyMonkey*.

4.2.1 A amostra

O estudo restringiu-se aos RDCs cadastrados no OpenDoar até a data de início do envio dos questionários, 4 de setembro de 2014. A necessidade de utilizar o *Survey* para coleta dessas informações surgiu após ter se mostrado inviável a obtenção de forma automática ou pela inconsistência de informações disponibilizadas pelo DSpace OAI-PMH *data provider*, quando comparadas com os *sites* dos RDCs.

Foram selecionados apenas repositórios que possuíam *e-mails* cadastrados, independentemente da tipologia e localização geográfica, o que correspondeu a 2.165, de um total de 2.720 repositórios digitais registrados no OpenDoar. Optou-se por utilizar o OpenDoar para obtenção de informações, como *e-mails* dos repositórios, por ser um diretório de repositórios digitais já utilizado como fonte de informação por projetos e iniciativas europeias, como o Digital Repository Infraestructure Vision for European Research – Diretrizes Driver, organizações como Surf Foundation, e por ser mantido pelo Sherpa Services – University of Nottingham (http://www.sherpa.ac.uk/), organização conceituada no contexto da comunicação científica e políticas de acesso aberto.

4.2.2 O questionário

A coleta de informação dos RDCs foi feita entre setembro de 2014 e novembro de 2014. O questionário, estruturado em três seções, com o total de 11 questões fechadas, foi respondido por 431 gestores de RDCs (http://repositorium.sdum.uminho.pt/handle/1822/1). A primeira seção visou à identificação do

repositório; a segunda, à verificação da existência de PAMs; e a terceira, à identificação dos esquemas e elementos de metadados utilizados pelos RDCs (APÊNDICE A).

4.2.3 O resultado

O resultado do *Survey* foi apresentado na International Conference on Dublin Core and Metadata Applications, em São Paulo, no ano de 2016 (http://hdl.handle.net/1822/37855), sendo aqui apresentada apenas uma síntese dos resultados.

A partir da identificação do nome ou acrônimo do repositório, foi solicitada a identificação do tipo de RDC (Questão n.º 2 – Q2). Neste estudo, entende-se que os RIs são responsáveis por armazenar a produção intelectual de uma instituição; RT estão vinculados a um domínio; repositório organizacional (RO) está associado a uma organização que, necessariamente, não segue a mesma finalidade dos institucionais de armazenar a produção científica; ROE armazena objetos educacionais; e o repositório de teses e dissertações (RTD) armazena teses e dissertações (Armbruster & Romary, 2010, Darby et al., 2008). Do total de 401 questões respondidas, 30 foram ignoradas. Entretanto, como essa questão possibilitava mais de uma resposta, 69 respondentes classificaram os seus repositórios em mais de uma opção, RI e ROE, por exemplo. Desse total, dez indicaram que são RO e RI, o que, em nosso entender, retrata uma contradição em razão das definições supracitadas. Assim, buscou-se confirmar no http://www.opendoar.org/ como esses dez repositórios estão categorizados e adicionou-se essa informação aos tipos correspondentes. Quanto aos outros tipos não houve a necessidade de fazer esse tipo de avaliação por entender que não existe conflito terminológico, podendo existir mais de um tipo de repositório em uma instituição. Nesse sentido, considerou-se o índice apontado pelos respondentes, em que identificou-se que os RIs são predominantes, 358; seguidos pelos RTs,36; ROs,25; RTDs, 52; e ROEs, 15.

Os artigos científicos figuram, com base nos respondentes, como o tipo de recurso mais armazenado, seguidos por livros/capítulos e teses e dissertações. Foi citado ainda pelos respondentes o armazenamento de: conjunto de dados, aparições em mídias, documentos administrativos e técnicos, *blogs* acadêmicos, programas curriculares e literatura cinzenta (Figura 17). Adicionalmente foi referido o uso de metainformação, ou seja, disponibilizam-se apenas os atributos que descrevem um recurso, sem o conteúdo. Para alguns autores, é questionável, se neste caso pode se denominar RDC.[28] Considerando

[28] Opiniões obtidas de modo informal em eventos da área.

os tipos de repositórios e os tipos de recursos armazenados, percebe-se que os repositórios vêm armazenando recursos variados, independentemente da sua tipologia, corroborando as observações de (Heery, 2009, p.13).

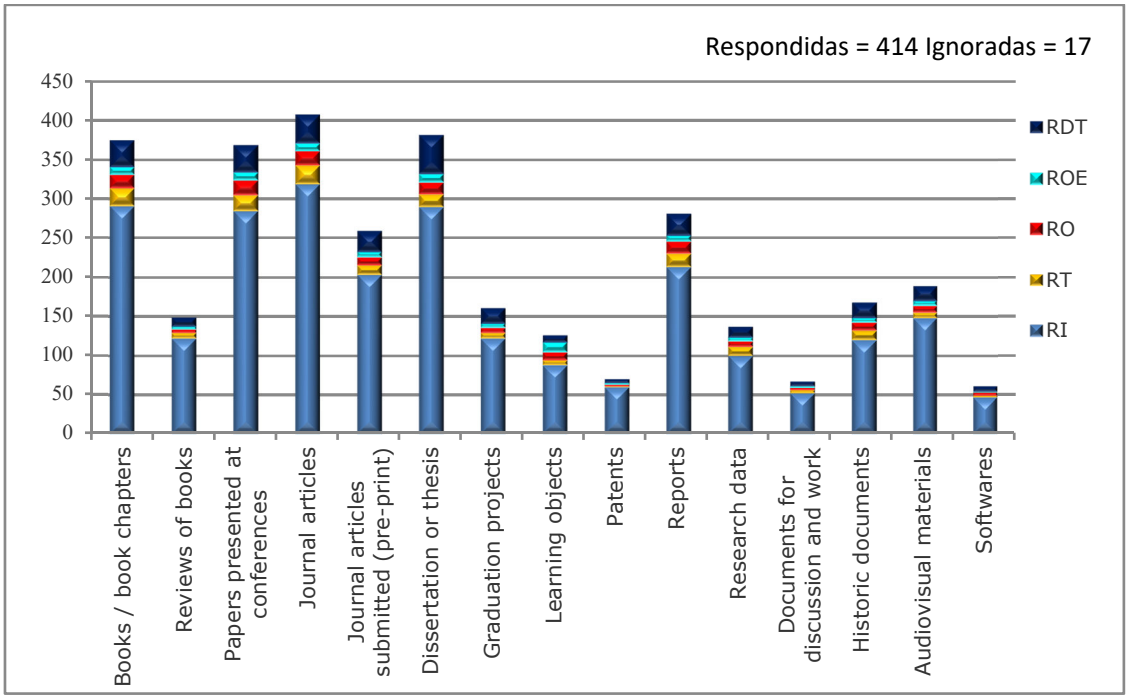


Figura 17. Tipos de materiais armazenados nos RDCs
Fonte: Andrade & Baptista (2015).

Investigou-se também se os RDCs adotavam recomendações internacionais que objetivam o aumento da qualidade dos metadados desses repositórios. Alguns respondentes afirmam que utilizam as recomendações da DCMI (http://dublincore.org/documents/dces/; http://dublincore.org/documents/2012/06/14/dcmi-terms/; http://dublincore.org/documents/profile-guidelines/) e o uso de protocolo OAI-PMH (http://www.openarchives.org/pmh/). No que concerne ao total de resposta, um número relativamente baixo de repositórios refere adotar o SKOS (http://www.w3.org/2004/02/skos/) e a OWL (http://www.w3.org/2001/sw/wiki/OWL). Nota-se, entretanto, um número maior dos que declaram utilizar especificações RDF (http://www.w3.org/standards/techs/rdf#w3c_all; http://www.w3.org/RDF/). O padrão OAI-ORE (http://www.openarchives.org/ore/) tem nos RIs os seus maiores adeptos. Foi citado ainda a utilização do DRIVER.

A literatura é clara em relação à utilização de recomendações, como as da DCMI e o uso de protocolo OAI-PMH, como forma de garantir a interoperabilidade entre os repositórios. As Diretrizes Driver, assim como Open Archives Initiative Object Reuse and Exchange (OAI-ORE), têm grande alcance na comunidade de RDCs porque norteiam o estabelecimento de condutas para o seu desenvolvimento. A utilização do

SKOS e a OWL contribui para a interoperabilidade semântica, enquanto a adoção das RDFs viabiliza, sobretudo, a ligação de dados (*Linked Data*) entre diferentes sistemas. Assim sendo, essas recomendações deveriam ser adotadas como forma de contribuir para o aprimoramento dos repositórios em relação aos seus propósitos, como exposto no início deste estudo.

Observou-se, pela respostas dos gestores, que o número de repositórios que possuem PAMs ainda é muito baixo. Diante desse resultado, podem ser consideradas duas opções para os RDCs não possuírem PAM: o desconhecimento do que seja um PAM ou o fato de as pessoas[29] desenvolverem o conteúdo referente ao perfil sem saber que se trata de um PAM. Esse fato também pode estar relacionado com as diferentes terminologias adotadas para o PAM ("*application profile*", "*metadata schema*", "*metadata vocabulary*" e "*dictionary metadata*"), o que dificulta a descoberta de RDCs que fazem uso desse tipo de aplicação (Andrade & Baptista, 2015).

No que se refere aos esquemas de metadados adotados, a predominância é do Dublin Core Metadata Element Set (informado pelos respondentes às vezes como DC Simples e Qualified - http://wiki.dublincore.org/index.php/User_Guide), seguido por Open Archives Initiative-Dublin Core (OAI-DC), que representa o uso dos elementos DC utilizados pela iniciativa do acesso aberto (OAI), mas que alguns gestores identificam como coisas distintas; METS e Machine-Readable Cataloguing (MARC); Metadata Object Description Schema (MODS); Electronic Thesis and Dissertation Metadata Standard (ETD-MS); LOM; DSpace intermediate metadata (DIM); Multimedia Content Description Interface (MPEG21); e Academic Metadata Format (AMF) (Figura 18). Aqui vale a pena esclarecer que as pessoas ainda apelidam de DC qualified (expressão caída em desuso na comunidade DCMI) o conjunto de elementos DC mais os seus elementos de refinamentos. É preciso salientar que a plataforma DSpace inclui no chamado "DC Qualified" elementos de metadados não pertencentes ao DC Terms e que foram incluídos no âmbito do desenvolvimento dessa plataforma.

[29] Declarações de participantes do Workshop sobre Perfil de Aplicação ministrado durante a Dublin Core International Conference, São Paulo, 2015.

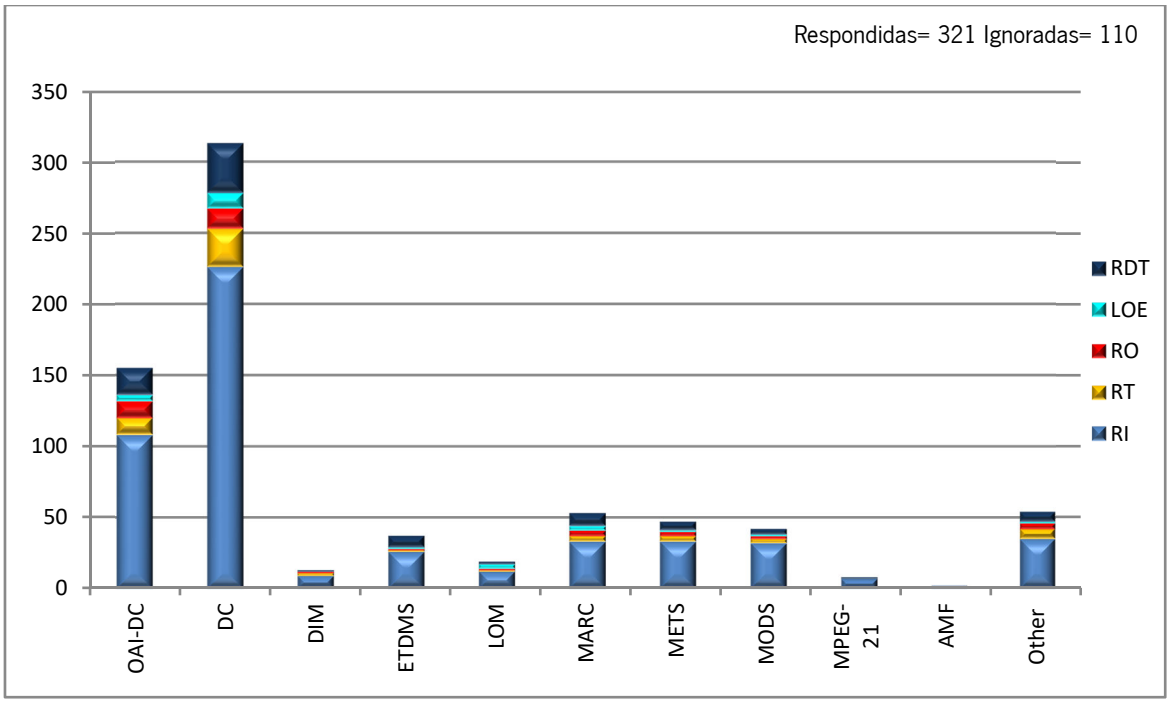


Figura 18. Esquema de metadados utilizados por RDCs.
Fonte: Andrade & Baptista (2015).

As respostas dos gestores em relação ao uso de metadados mostram ainda que: a) alguns repositórios fazem pouco uso de esquemas de metadados desenvolvidos especificamente para descrição de determinados tipos de recursos, prevalecendo o uso do DC. Como exemplo, pode-se citar o caso dos ROEs e de RTDs que, embora tenham sido projetados para armazenarem objetos educacionais, teses e dissertações, respectivamente, utilizam predominantemente o DC; b) repositórios que armazenam diferentes tipos de recursos se restringem ao uso do DC, não explorando outros esquemas que poderiam enriquecer a descrição dos recursos; e c) na utilização de diferentes esquemas e elementos de metadados por determinados repositórios, seria aconselhável a definição de PAM, o que clarificaria para humanos e máquinas informações sobre os metadados utilizados e potencializaria a interoperabilidade de metadados.

Diante desse resultado, é possível sugerir que a predominância do uso do DC pode ser justificada pelo fato de os RDCs ainda trabalharem com o nível de interoperabilidade básico proporcionado pelo uso do DC (OAI-DC), em que a recolha de dados é baseada no protocolo OAI-PMH. A utilização de esquemas de metadados projetados para uma descrição mais granular e específica de um determinado recurso possibilita um avanço em termos de interoperabilidade semântica, ainda mais se estiverem em conformidade com os princípios *Linked Data* (c.f. Seção 2.1.1 *Linked Data*). Organizações, como a própria DCMI e W3C, disponibilizam recomendações para que, ao se utilizar diferentes esquemas de

metadados, possam proporcionar a otimização da troca de informações entre os diversos serviços de informação (http://dublincore.org/documents/dces; http://dublincore.org/documents/profile-guidelines/; http://dublincore.org/documents/dc-dsp/; http://dublincore.org/documents/singapore-framework/; http://www.w3.org/2004/02/skos/; http://www.w3.org/2001/sw/wiki/OWL; http://www.w3.org/RDF/; http://dublincore.org/documents/dsp-wiki-syntax/; http://www.openarchives.org/ore/). Recentemente foi publicado a Shapes Constraint Language (SHACL) Primer, uma linguagem para validação de gráficos RDF que pode ser entendida como uma descrição dos gráficos de dados que satisfazem condições de um gráfico RDF. Essas descrições podem ser utilizadas para construção de interface de usuário, geração de código e integração de dados (Knublauch, TopQuadrant, Inc., Prud'hommeaux, & W3C/MIT, 2015, Knublauch & Kontokostas, 2017). No contexto da comunicação científica, por exemplo, essa linguagem pode contribuir para ordenação de autores de um artigo, não pela ordem alfabética, mas por razões específicas como atribuição de mérito na elaboração do artigo (Steyskal & Coyle, 2017)

Em uma segunda fase da análise do uso de esquemas de metadados, foram identificados os elementos utilizados pelos RDCs. Como descrito, o esquema mais adotado é o DC. Os cinco elementos mais usados foram: *title*, *author*, *description*, *date* e *type*. Os respondentes indicaram ainda o uso dos seguintes elementos: *alternative*, *bibliographicCitation*, *isPartOf* e *audience*. Também foram informados elementos que não fazem parte do DCMI Terms, que podem ter sido adicionados no âmbito do DSpace: *placeOfPublication*, *root*, *series*, *number*, *edition*, *volume*; *level of audience*; *dc.contributor.author*, *dc.subject.other*, *author contact*, *editor contact*, *date available*, *date accessioned*, *start page*, *end page*, *ispartofname*, *ispartofnumber*, *ispartoftitle* e *ispartofvolume*.

Quanto ao esquema de metadados LOM, as classes mais utilizadas, segundo os respondentes, foram a General e a Educational e os elementos: *General -> description*; *General -> identifier -> catalog*; *General -> title*; *General -> language* e *Educational -> learning resource type*. Verificou-se ainda a baixa utilização do LOM para descrição dos objetos educacionais, até mesmo pelos ROEs.

Com base nas informações dos respondentes, os elementos do MODS foram adotados pelo menos uma vez, sendo utilizados em maior parte pelos RIs. Os cinco mais utilizados foram: *titleInfo -> title*; *originInfo -> Publisher*, *abstract* e *originInfo -> dateIssued*; e *subject -> topic*.

Outro esquema bastante adotado foi o MARC 21. Pode-se justificar esse resultado pela própria simplicidade do DC e pela história do MARC 21 no domínio dos serviços de informação

Verificou-se que os RIs são o tipo de repositório que mais explorou o uso de diferentes esquemas de metadados. Entretanto, percebe-se que, embora os demais acompanhem a tendência dos RIs, os ROEs e RTDs não exploram tanto os esquemas de metadados que foram desenvolvidos para tipos específicos de recursos, como pode ser observado na Tabela 37.

Tabela 37. **Esquemas de metadados adotados por RDCs**

Esquemas de Metadados	Tipos de repositórios				
	RI	**RT**	**RO**	**ROE**	**RTD**
AMF	2	0	0	0	0
DC/OAI-DC	334	39	26	15	54
DIM	9	2	1	0	1
ETDMS	26	1	1	1	8
LOM	12	1	1	3	2
MARC	33	4	4	3	9
METS	33	4	3	1	6
MODS	32	3	2	1	4
MPEG-21	7	0	0	0	1

Os RDCs analisados que adotam PAMs, independentemente da sua tipologia, contrastam com as recomendações da DCMI que estabelecem o uso dos PAMs para a otimização da interoperabilidade semântica (Coyle & Baker, 2009). O pouco conhecimento, por parte dos administradores, sobre PAMs e suas vantagens pode ser a razão do baixo número de PAMs adotados pelos repositórios RDCs.

4.3 BASES DE DADOS BIBLIOGRÁFICAS NO DOMÍNIO DA ONCOLOGIA

A identificação dos campos de busca adotados por bases de dados bibliográficas preenche uma lacuna na literatura que é a identificação de campos potencialmente associáveis aos metadados descritivos. Nesse sentido, foram analisadas bases de dados bibliográficas que contemplam o domínio da Oncologia em relação aos seguintes aspectos: analisar a interface de busca simples e avançada para identificação e registro dos pontos de acesso (elementos de metadados); detectar e registrar refinamentos e facetas utilizadas; comparar os pontos identificados; elencar o conjunto de pontos com potencial para o aumento da qualidade dos resultados de investigação no contexto dos RDCs.

Essas questões visam a revelar elementos disponibilizados pelas bases de dados que possam estar alinhados às NIs de oncologistas e identificar tendências em relação aos pontos de acesso.

4.3.1 A amostra

A amostra se restringiu às bases de dados disponibilizadas pela *b-on* (Universidade do Minho) e pelo portal dos Periódicos da Coordenação de Aperfeiçoamento de Pessoal Nível Superior – CAPES (Tabela 38).

Tabela 38. **Bases de dados bibliográficas analisadas**

Base de Dados	Tipo
SCOPUS (Elsevier)	Referencial
Web of Science (Thomson Reuters Scientific)	Referencial
Annual Reviews	Texto completo
MEDLINE 1946 to present with Weekly Update (Ovid)*	Referencial com resumos
MEDLINE completa (EBSCO)*	Referencial – Texto completo
MEDLINE/PubMed (National Library of Medicine)*	Referencial com resumos
Bentham Science	Texto completo
Cambridge Journals Online	Texto completo
Science Direct (Elsevier)	Texto completo
Wiley Online Library	Texto completo
Gale Academic One File	Texto completo
Academic Search Premier	Referencial - Texto completo
HighWire Press	Texto completo

Nota: A Medline é uma base de dados explorada por diferentes editoras que adotam customização distintas, por essa razão foi considera cada base (plataforma) individualmente.

4.3.2 A coleta de dados

Para a coleta de dados foi elaborado um formulário em Word, com base no de estudo de Falagas, Pitsouni, Malietzis, & Pappas (2008) e análise prévia das BDBs. Foram definidos campos para caracterização da BDB e identificação dos pontos de acesso disponíveis, tanto para a busca simples, quanto a avançada, e opções de refinamento de buscas. A planilha foi utilizada para todas as bases, na medida que novos pontos de acesso fossem surgindo, eles eram acrescentados ou modificados (Tabela 39).

Tabela 39. **Formulário para análise das bases de dados bibliográficas**

Identificação : Nome da base de dados				
Proprietário				
Data de criação				
Descrição				
Idioma				
Período de Cobertura				
Número de Revistas				
Foco (área)				
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos				
Links				
Resumo				
Gerenciamento biblioteca				
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract				
Article title				
Author				
Author name				
Caption				
Cited Author				
Country				
Document type				
DOI				
Full text				
Group author				
ISBN				
ISSN				
Keywords				
Language				
Publication date				
Publication title				
Publication year				
Source Title				
Source Type				
Subject Area				
Topic				
Year				

Os dados coletados nas 13 bases de dados estão disponibilizados nas Tabelas 40 a 52.

Tabela 40. **SCOPUS**

1**) IDENTIFICAÇÃO – SCOPUS**				
Proprietário	Elsevier			
Data de criação	2004			
Descrição	Referenciais com resumos, multidisciplinar, indexa patentes e fontes de informação na *WEB*			
Idioma	Inglês			
Período de Cobertura				
Número de Revistas	21 mil periódicos 24 milhões de patentes 5 mil editores internacionais			
Foco (área)	Ciências Biológicas, Multidisciplinar, Ciências Exatas e da Terra, Ciências Ambientais, Ciências da Saúde, Ciências Agrárias, Engenharias, Ciências Sociais Aplicadas, Ciências Humanas, Linguística, Letras e Artes			
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos	Serviço de alerta,APIs, compara índices de journals; Consulta a Índice H dos autores, SJR, SNIP e IPP dos periódicos ; Fornece indicadores altmétricos do documento			
Links	http://www.scopus.com/			
Resumo				
Gerenciamento biblioteca				
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract	X	x	Áreas	
Address			Período	
Affiliation	X	x	Tipo de documento	
Article title	X	x		
Author identifiers				
Author name	X	x		
Caption				
Cited Author				
Country				
Document type				
DOI	X	x		
Full texto				
Group author				
ISBN				

1) **IDENTIFICAÇÃO – SCOPUS**				
ISSN	X	x		
TIPOS DE PESQUISA	campos de busca		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Keywords	X	x		
Language	X	x		
Publication date				
Publication title				
Publication year				
Source Title				
Source Type				
Subject Area				
Topic				
Year				
First author	X	x		
Source title	X	x		
Affliation name	X	x		
Affiliation city	X	x		
Affiliation country	X	x		
CODEN	X	x		
References	X	x		
Conference	X	x		
Chemical name	X	x		
CAS Number	X	x		

Tabela 41. **Web of Science**

2 **) IDENTIFICAÇÃO - Web of Science**	
Proprietário	Thomson Reuters Scientific
Data de criação	2004
Descrição	Referênciais com resumos, texto completo, indexa periódicos, livros e anais
Idioma	Inglês. Interface em vários idiomas. Algumas bases em coreano e chinês
Período de Cobertura	1945-
Número de Revistas	12.000 research journal (inclui Open Access Journal)
Foco (campos/ área)	Ciências Biológicas, Multidisciplinar, Ciências Exatas e da Terra, Ciências Ambientais, Ciências da Saúde, Ciências Agrárias, Engenharias, Ciências Sociais Aplicadas, Ciências Humanas, Linguística, Letras e Artes
Bases de dados cobertas	Science Citation Index a partir de 1945; Social Sciences Citation Index a partir de 1956; Arts and Humanities Citation Index a partir de 1975; a partir de Conference Proceedings Citation Index-Science, Conference Proceedings Citation Index-Social Science e Humanities; Emerging Souces Citation Index, a partir de 2015; Derwent innovation índex a partir de 1963; Biological Abstraccts a partir de 1926; Korean Journal Database, Scielo Citation Index. Book Citation Index, Current Chemical Reactions; Index chemicus; a partir de 1926:; Biosis Preview Science, Biosis Citation Index, CAB Abstract a partir de 1910; INSPEC a partir de 1898, MEDLINE a partir de 1950, Current Contents Connect, SCI Data Citation Index, Chinese Science Citation Database; a partir de 2012: FSTA Food Science and Technology Abstract a partir de 1969; Zoological Record a partir de 1864.
Número de palavras-chave permitidas	6.000 termos e três campos de busca
Usos	Itens relacionados; ferramentas de análises; busca por autor; busca por referência citada (autor, publicação e ano citados), verificador ortográfico na digitação dos termos
Links	www.webofscience.com
Ferramentas gestão	EndNote - gerenciamento de bibliotecas ResearchID – gerenciamento de identificador para autores InCites – ferramenta para avaliação de citações dos autores e periódicos e instituições e indicadores de pesquisa Serviço de alerta e RSS Feeds Relatório de citações Relatório de uso

ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Topic	X	x		Article
Title	X	x		Abstract of published item
Author	X	x		Art Exhibit Review
Author name	X	x		Bibliography
Group author	X	x		Biogrphical-Item
Editor	X	x		Book
Publication title	X	x		Book Chapter

2) **IDENTIFICAÇÃO - Web of Science**				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
DOI	X	x		Book Review
Publication date	X	x		Chronology
Address	X	x		Correction
Organização – Consolidada	X	x		Correction, Addition
Conference	X	x		Dance performance review
Language	X			Database review
Document type	X			Discussion
Agency funding	X	x		Editorial material
Funding name	X	x		Excerpt
Access number	X	x		Fiction, Creative prose
ID PubMed	X	x		Film review
Organization		x		Hardware review
Suborganization		x		Item about na individual
Name of street		x		Letter
City		x		Meeting abstract
Province/State		x		Meeting summary
Country		x		Music performance review
Postal code		x		Music score
Text of funding		x		Music score review
Search field				News item
Category *Web* of Scielo		x		Note
ISSN/ISBN		x		Poetry
				Proceedings paper
				Record review
				Review
				Script
				Software review
				TV Review
				TV Review, Radio Review
				TV Review, Radio Review, Video
				Theater review

Tabela 42. **Annual Reviews**

IDENTIFICAÇÃO - Annual Reviews				
Editor	Annual Reviews			
Data de criação	1932			
Link				
Descrição	Textos completos. Artigos de revisão de literatura.			
Idioma	Inglês			
Período de Cobertura				
Número de Revistas	46			
Foco (área)	Multidisciplinar. Biomedical, Life, Physical, Social Sciences			
Bases de dados cobertas	49 periódicos			
Número de palavras-chave permitidas				
Usos	Descrição dos metadados utilizados, salva os resultados das pesquisas, artigos relacionados			
Links				
Resumo				
Gerenciamento biblioteca	RSS com alerta sobre inclusão de novos metadados			
Citação	CABAbstracts, Chemical Abstracts Service, Compendex, EconLit, EMBASE, FLUIDEX, GEOBASE, MEDLINE/PubMed, ProQuest/CSA, Scopus,, Social Science Research Network, Thomson Reuters /ISI *Web* of Knowledge			
Dicas	Busca por autor: primeiro e ultimo nome			
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x	Título de periódico	Data de publicação
Keywords		x		
Language				
Publication date				
Publication title				
Publication year				
Source Title				
Source Type				
Subject Area				
Topic				
Year				
Article title				
Authors		x		
Cited Author		x		
Full text		x		
Captions		x		
Terms	X			
Journals	X			

Tabela 43. **MEDLINE 1946 to Present with Weekly Update (Ovid)**

IDENTIFICAÇÃO- MEDLINE 1946 to Present with Weekly Update (Ovid)	
Proprietário	Ovid Technologies
Data de criação	
Link	
Descrição	Referenciais com resumos
Idioma	Interface em vários idiomas
Período de Cobertura	
Número de Revistas	4.600 periódicos
Foco (área)	Ciências Biológicas, Ciências da Saúde, Ciências Agrárias,
Bases de dados cobertas	MedLine, Healthstar
Número de palavras-chave permitidas	Busca por assunto procura um termo exato no assunto inteiro (Assunto MeSH).
Usos	Termos relacionados, inclusão de multimídia, armazenamento em biblioteca, exporta para gerenciadores de bibliotecas e outros formatos
Links	
Resumo	
Gerenciamento biblioteca	
Citação	

ESTRUTURA DAS BASES DE DADOS

TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x	x	
Abstract label		x		
Address				
Affiliation				
Article title				
Author identifiers				
Category		x		
Comments		x		
Country of publication		x		
Date Created		x		
Date of publication		x		
DOI		x		
Electronic date of publication		x		
Grant number		x		
Group author				
Humans			x	
Institution		x		
Investigator		x		
Investigator nameID		x		
ISSN linking		x		

IDENTIFICAÇÃO- MEDLINE 1946 to Present with Weekly Update (Ovid)				
Issue/Part		x		
Journal abbreviation		x		
Journal name		x		
Journal subset		x		
Journal Word		x		
Keyword heading		x		
Keyword heading Word		x		
Keywords		x		
Language		x		
Latest Update			x	
Media type		x		
MeSH Subject Heading		x		
Name of substance Word		x		
NLM Category		x		
NLM Journal code		x		
NLM Journal name		x		
NLM Journal Word		x		
Original title		x		
Other ID		x		
Page		x		
Personal name as subject		x		
Pharmacologic actions			x	
Primary author		x		
Protocol supplementary		x		
Publication date				
Publication title	**x**			
Publication type		x		
Publication year				
Publisher item identifier		x		
Rare disease supplementary		x		
Record owner		x		
Registry number/name of subs		x		
Review Articles			x	
Revision date		x		
Secondary source ID		x		
Secondary source NA		x		
Secondary source*/link*		x		
Source				
Space flight mission		x		
Status		x		
Structured abstracts			x	
Subject heading word		x		
Text Word		x		
Title		x		

IDENTIFICAÇÃO- MEDLINE 1946 to Present with Weekly Update (Ovid)				
Unique identifier		x		
Update date		x		
Version date		x		
Version ID		x		
Volume		x		
Year of publication		x		
Abstract		x	x	
Abstract label		x		
Address				
Affiliation				
Article title				
Author identifiers				
Author last name		x		
Author name	x			
Author nameID		x		
Authors		x		
Authors full name		x		
Category		x		
Comments		x		
Copyright índex		x		
Core Clinical Journals			x	
Country of publication		x		
Date Created		x		
Date of publication		x		
DOI		x		

Tabela 44. **MEDLINE Complete (EBSCO)**

IDENTIFICAÇÃO- MEDLINE Complete (EBSCO)				
Editor	EBSCO			
Data de criação				
Link	http://www-epnet-com.ez43.periodicos.capes.gov.br/			
Descrição	Referenciais com resumos e textos completos			
Idioma	Inglês			
Período de Cobertura	1949			
Número de Revistas	4.800 títulos de periódicos. 2.100 periódicos com texto completo a partir de 1965			
Foco (área)	medicina, enfermagem, odontologia, medicina veterinária, saúde pública, pré-clínicos. Ciências Humanas, Ciências Agrárias, Ciências Biológicas, Ciências da Saúde			
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos	Pesquisa SmartText (frases, sentença, parágrafo), palavras relacionadas, pesquisa em texto completo, assuntos relacionados			
Links				
Resumo				
Gerenciamento biblioteca				
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x		
Address				
Affiliation				
Article title				
Author identifiers				
Author name		x		
Caption				
Cited Author				
Country				
Document type				
DOI				
Full text				X
Group author				
ISBN		x		
Subject Area				
Topic				
Year				
ISSN		x		

IDENTIFICAÇÃO- MEDLINE Complete (EBSCO)				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA		PESQUISA SIMPLES
Keywords				
Language				
Publication date				
Publication title				
Publication year				
Source Title		x		
Source Type				
Visualização de imagem	X			
Revistas acadêmicas	X			
Referências disponíveis	X			
publicação	X			
sexo	X			
Citações	X			
Title		x		
Tipo de publicação		x		
Humano ou animal		x		
Consultas clinicas		x		
Idioma		x		

Tabela 45. **PubMed (via National Library of Medicine)**

IDENTIFICAÇÃO – PubMed (via National Library of Medicine)				
Editor	United States National Library of Medicine (NLM)			
Data de criação	1997			
Link	www.nlm.nih.gov			
Descrição	Referenciais com *link* para texto completo (PubMed Central e outros *sites*). Texto completo de periódicos OA. Se o acesso é institucional, possibilita acesso direto a periódicos.			
Idioma	Inglês			
Período de Cobertura	1940			
Número de revistas	25.000 milhões de citações (artigos e livros)			
Foco (área)	Ciências Biológicas, Ciências Agrárias, Ciências da Saúde, biomedicine and health, broadly defined to encompass those areas of the life sciences, behavioral sciences, chemical sciences, and bioengineering, biology, environmental science, marine biology, plant and animal science as well as biophysics and chemistry			
Bases de dados cobertas	MedLine, Journal Instituto Nacional de Câncer dos Estados Unidos (NCBI) Databases			
Número de palavras-chave permitidas				
Usos	Utilização do MESH Terms, publication types, GenBank accession numbers, Clinical queries, Topic-Specifiq Queries, RSS, Alerta			
Links				
Abstract				
Gerenciamento biblioteca	PubMed Bibliography			
Citação	Single Citation Matches			
Vocabulário Controlado	MeSH			
Formato final	Summary, Summary (text), Abstract, Abstract (text), MEDLINE, XML, PMID List			
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x		Tipos de artigos
Address		x		Disponibilidade de texto (abstract, Free Full Text, Full Text)
Affiliation		x		Trending articles
Article title		x		Publication dates
Author identifiers		x		Species
Author Corporate		x		Language
Author First		x		Sex
Author Full		x		Subject
Author Last		x		Journal categories
Book		x		Age
Caption		x		Search field
Cited Author				

IDENTIFICAÇÃO – PubMed (via National Library of Medicine)				
Volume		x		
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Date MeSH		x		
Date Modification		x		
Date Publication		x		
EC/RN Number		x		
Country		x		
Document type		x		
DOI		x		
Editor		x		
Filter		x		
Grant Number		x		
Investigator		x		
Investigator Full		x		
ISBN		x		
Issue		x		
Journal		x		
Location ID		x		
Language		x		
Location ID		x		
MeSH Major Topic		x		
MeSH Subheading		x		
MeSH Term		x		
Other Term		x		
Pagination		x		
Pharmacological Action		x		
Publication Type		x		
Publisher		x		
Publication year		x		
Secondary Source ID		x		
Subject – Personal Name		x		
Transliterated title		x		
Supplementary Concept		x		
Text Word		x		
Title		x		
Title/Abstract		x		
Transiliterated Title		x		

Tabela 46. **Bentham Science Publishers**

IDENTIFICAÇÃO -Bentham Science Publishers				
Editor	Ingentaconnect			
Data de criação				
Link				
Descrição	Texto completo			
Idioma				
Período de Cobertura				
Número de Revistas	161 journals, 4 print/online book (series)			
Foco (área)	Pharmaceutical, biomedical and medical research community			
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos				
Links				
Abstract				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract				
Address				
Affiliation				
Article title		x		
Author identifiers				
Author name				
Document type				
DOI		x		
Full texto				
Group author				
ISBN		x		
ISSN		x		
Keywords				
Language				
Publication date				
Publication title	x			
Source Title				
Source Type				
Subject Area				
Topic				
Year				
Publisher	x			
Subject	x			
Article title, keywords or abstract		x		
Article title				
Article fulltex		x		

Tabela 47. **Cambridge Journals Online**

IDENTIFICAÇÃO - Cambridge Journals Online				
Proprietário	Cambridge University Press			
Data de criação				
Link	journals.cambridge.org			
Descrição				
Idioma	Ingles			
Período de Cobertura				
Número de Revistas	300			
Foco (área)	Multidisciplinar			
Bases de dados cobertas	Journals e livros publicados pela editora			
Número de palavras-chave permitidas				
Usos				
Links				
Resumo				
Gerenciamento biblioteca				
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x		
Address				
Affiliation		x		
Article title				
Author identifiers				
Author name				
Caption				
Cited Author				
Country				
Document type				
DOI		x		
Full texto		x		
Group author				
ISBN		x		
ISSN		x		
Keywords		x		
Language				
Publication date		x		
Publication title				
Source Title				
Source Type				
Subject Area				
Topic				
Year				
Article or chapter title				x

IDENTIFICAÇÃO - Cambridge Journals Online				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Author or Editor				x
EISSN				
Article subtitle				x
Author				x
Journal title				x
Volume nº				x
Issue nº				x
Issue date				x
Page range				x
Date				x

Tabela 48. **Highwire Press**

IDENTIFICAÇÃO - Highwire Press				
Editor	Stanford University			
Data de criação				
Link	highwire.stanford.edu			
Descrição	Texto completo			
Idioma	Inglês			
Período de Cobertura	1753-2016			
Número de Jornais				
Foco (área)	Ciências biológicas e Saúde			
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos	Gerenciador de bibliografias, RSS			
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract				
Affiliation				
Article title				
Author identifiers				
Author name		x		
DOI				
Full texto		x		
Subject Area				
Topic	x			
Year				
Title and Abstract only		x		
Title		x		
Publisher	X			

Tabela 49. **Science Direct**

IDENTIFICAÇÃO - Science Direct				
Editor	Elsevier			
Data de criação				
Link	www.sciencedirect.com			
Descrição	Textos completos			
Idioma	Inglês			
Período de Cobertura	1823-2016			
Número de Revistas	3.500 jornais e 34 mil livros			
Foco (área)	Ciências Biológicas, Ciências da Saúde, Ciências Agrárias, Ciências Exatas e da Terra, Engenharias, Ciências Sociais Aplicadas, Ciências Humanas e Letras e Artes			
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos	Inclui imagens e referências			
Links				
Resumo				
Gerenciamento biblioteca				
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x		
Address				
Affiliation		x		
Article title				
Author identifiers				
Author name		x		
DOI				
Full text				
Group author				
ISBN		x		
ISSN		x		
Keywords		x		
Language				
Publication date				
Publication title				
Source Title		x		
Subject Area				x
Topic				
Abstract, Title, Keywords				
Specific Author		x		
Title		x		
References		x		
Tipo de publicação				x

Tabela 50. **Wiley Online**

IDENTIFICAÇÃO – Wiley Online Library				
Editor	Wiley InterScience			
Data de criação	1999			
Link	www.onlinelibrary.wiley.com			
Descrição	Texto completo			
Idioma	Inglês			
Período de Cobertura				
Número de Revistas	2.375 periódicos 19.140 livros			
Foco (área)	Ciências Biológicas, Ciências da Saúde, Ciências Exatas e da Terra, Ciências Agrárias, Ciências Sociais Aplicadas, Ciências Humanas e Linguística, Letras e Arte			
Bases de dados cobertas				
Número de palavras-chave permitidas				
Usos				
Citação				
ESTRUTURA DAS BASES DE DADOS				
TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x		
Address				
Affiliation		x		
Article title		x		
Author identifiers				
Author name		x		
Caption				
Cited Author				
Document type				
DOI		x		
Full text		x		
Group author				
ISBN		x		
ISSN		x		
Keywords		x		
Language				
Publication date				
Publication title		x		
Publication year				
Source Title				
Subject Area				
Topic				
Funding Agency				
References		x		

Tabela 51. **Gale Academic OneFile**

IDENTIFICAÇÃO –Academic Search Premier	
Editor	Gale Group
Data de criação	
Link	
Descrição	Textos completos
Idioma	
Período de Cobertura	
Número de Revistas	
Foco (área)	Ciências Exatas, Tecnologia, Ciências Médicas, Ciências Sociais Aplicadas, Artes e Humanidades
Bases de dados cobertas	multidisciplinar – podem ser escolhidas por assunto
Número de palavras-chave permitidas	
Usos	pesquisa por palavras e assuntos relacionados
Links	
Resumo	
Gerenciamento biblioteca	
Citação	

ESTRUTURA DAS BASES DE DADOS

TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract		x		
Address				
Article title				
Author identifiers				
Author name		x		x
DOI				
Full texto	x	x		
Group author				
Document type				x
ISBN		x		
ISSN		x		
Keywords				
Language				
Publication date				
References				
Publication title				x
Publication subject				x
Tipo de visualização imagem				
Title				
Subject	x			
Publication	x			

Tabela 52. **Academic Search Premier**

IDENTIFICAÇÃO –Academic Search Premier	
Editor	EBSCO
Data de criação	
Link	
Descrição	referenciais com resumos
Idioma	
Período de Cobertura	
Número de Revistas	2.500 journals full text, referências e resumos de artigos de 8.450 títulos
Foco (área)	
Bases de dados cobertas	multidisciplinar – podem ser escolhidas por assunto
Número de palavras-chave permitidas	
Usos	pesquisa por palavras e assuntos relacionados
Links	
Resumo	
Gerenciamento biblioteca	
Citação	

ESTRUTURA DAS BASES DE DADOS

TIPOS DE PESQUISA	CAMPOS DE BUSCA		LIMITES DE BUSCA	
PONTOS DE ACESSO	PESQUISA SIMPLES	PESQUISA AVANÇADA	PESQUISA SIMPLES	PESQUISA AVANÇADA
Abstract				
Address				
Article title				
Author identifiers				
Author name		x		x
DOI				
Full texto		x		x
Group author				
ISBN				
ISSN		x		
Keywords		x		
Language				
Publication date				
References	x			
Publication type	x			
Publication data	x			
Tipo de visualização imagem	x			
Title		x		
Subject		x		
Abstract or Author –Supplied Abstract		x		
Author-supplied keywords		x		

IDENTIFICAÇÃO –Academic Search Premier				
Geographic terms		x		
People		x		
Reviews & products		x		
DUNS number		x		
Ticker symbol		x		
Journal name		x		
Acession number		x		

4.3.3 O resultado

A análise das BDBs proporcionou a identificação de pontos de acesso (campos e limites de busca) (Tabelas 40-52) que foram classificados em uma planilha de acordo com as seguintes categorias: identificadores, autoria, data, afiliação, *abstract*, suporte financeiro, assunto, tipo de recurso, descrição do recurso e endereço. A atividade de classificação incluiu a verificação de ambiguidade/redundância, que, ao ser constatada, um dos pontos de acesso era eliminado (Tabela 53).

Tabela 53. **Categorização dos pontos de acesso identificados nas bases de dados bibliográficas**

Identificação	Autoria	Data	Afiliação	Agência de financiamento	Assunto	Tipo	Descrição	Cobertura	Refinamentos
CAS Number	Author Corporate	Date Completion	Affiliation city	Funding name	Category *Web* of Scielo	Conference	Volume	Address	Chemical name
CODEN	Author First	Date Create	Affiliation country	Funding agency	Floating sub-heading	Media type	Duration	Country	Citation
DOI	Author Full	Date Created	Institution		MeSH Major Topic	Book	Edition	Publication county	Cited Author
DUNS number	Author Last	Date Entrez	Organization		MeSH Subheading	Book chapter	Comments		Captions
EC/RN Number	Investigator	Date MeSH			MeSH Subject Heading	Publication data	Journal abbreviation		Exploded sub-heading
Grant number	Author name ID	Date Modification			MeSH Term		Journal subset		Fulltext
ID PubMed	First author	Version date			Personal name as subject		Journal categories		Gender
ISBN	Group author	Electronic date of publication			Subject		Language		Gene symbol
ISSN	Publisher	Period			Subject – Personal Name		Topic		Gene symbol Word
ISSN print	Cited author	Revision date			Subject heading word		Trending articles		Conflict of interest statement
EISSN		Update date			Supplementary Concept		Trasiliterated title		Human or animal
Unique identifier of investigator					Terms		Page		Pharmacological Action
Location ID					Author keyword		Pagination		Protocol supplementary
NAICS code or descrption					Keyword heading		Images		Rare disease supplementary

Identificação	Autoria	Data	Afiliação	Agência de financiamento	Assunto	Tipo	Descrição	Cobertura	Refinamentos
NLM Journal code					Keywords		Place name		References
Número de acesso					Search área		Publication title		Registry number/ name of subs
Publisher item identifier					Personal name as subject		Record owner		Reviews & products
Secondary Source ID					Name of substance		Abstract		Age
					Rare disease supplmentary		People		Race and gender
					Supplmentary concept		Brand name		Person-about
					Geographic terms		Source Type		Tipo de visualização imagem
							Space flight mission		Visualização de imagem
							Start Page		By line text
							Status		Reader comments
							Issue		Outline headings
							Issue/Part		Copyright index
							Article title		Reader comments
							References		Ticker symbol
							Secondary source		
							Secondary source link		

SÍNTESE DA SECÇÃO 4

Os estudos sobre o uso de perfil de aplicação por RDCs (Revisão de Literatura e *Survey*) e sobre as bases de dados bibliográficas indicam que:

a) em relação aos perfis de aplicação:

- seja por meio da revisão da literatura, seja pelo *Survey*, é visível que o número de RDCs que define PAM ainda é baixo, situação que pode ser justificada pelo desconhecimento dos gestores quanto a sua relevância para a interoperabilidade ou pelo uso de diferentes terminologias;
- os RIs são os que mais utilizam diferentes esquemas de metadados, característica que reforça o uso de PAM;
- a análise da literatura proporcionou a identificação de dois PAMs no domínio da Medicina que contemplam publicações: um para descrição de recursos disponibilizados na web (Governo do Canadá) e outro para biblioteca (NLM, Estados Unidos);
- a revisão de literatura e o *Survey* não possibilitaram a identificação de PAM específico para RDCs no domínio da Oncologia.

b) em relação aos esquemas e elementos de metadados

- o Dublin Core Element Set é o esquema de metadados mais utilizado pelos RDC investigados, justificado por características como simplicidade, flexibilidade e por ser a base do OAI-PMH; e
- alguns PAMs mostram que instituições desenvolvem seus próprios esquemas de metadados e/ou incluem elementos associados às extensões do DC.

c) em relação às bases de dados bibliográficas

- no geral as bases seguem o mesmo padrão de descrição, restringindo-se aos metadados descritivos, títulos – publicação e fonte –, autor, data e assunto. O ponto de acesso assunto ou tópico pode deixar de atender às NIs de utilizadores no momento em que não propicie a granularidade necessária à especificidades de cada domínio; e
- alguns pontos de acesso identificados em BDBs possuem correspondentes no DC. Nota-se que existem pontos de acesso mais adequados do que os elementos do DC para determinadas propriedades, por exemplo, os pontos de acesso da agência de financiamento.

Assim, foi possível identificar esquemas e elementos de metadados utilizados por RDCs e pontos de acesso adotados pelas BDBs, que podem vir a contribuir para o desenvolvimento de PAMs.

Diante do que foi analisado, observa-se que os PAMs identificados são predominantemente genéricos. A autora defende a relevância do desenvolvimento de perfis para repositórios voltados para domínios específicos, cujas especificidades inerentes a cada domínio deveriam ser observadas, o que contribuiria para a interoperabilidade semântica entre sistemas. Nesse sentido, a identificação das NIs, nomeadamente os TNIs, se torna relevante para todos os atores que atuam no contexto dos sistemas de informação, sobretudo para o utilizador-final.

5 ESTUDO DAS NECESSIDADES INFORMACIONAIS DOS ONCOLOGISTAS

> *As cancer research expands in both scope and complexity, management of data associated with cancer research becomes a technical and economic challenge. There would be clear merit in an application that could provide the majority of the information required by the majority of the cancer research projects across a given institution. Before such an application can be built, however, there must be a clear understanding of the information requirements of the cancer researchers in question*
>
> (Kim & Gilbertson, 2007, p. 1383)

Os estudos preliminares, apresentados na Seção 4, não indicaram a existência de trabalhos sobre PAMs para RDCs com base em TNIs no domínio da Oncologia. A partir dessa constatação, foi preciso desenvolver uma investigação, cujo objetivo foi identificar os TNIs dos profissionais (médicos e investigadores) que atuam na Oncologia. A escolha do domínio foi feita por considerar que a multidisciplinariedade, o volume e a heterogeneidade de informação podem comprometer a interoperabilidade semântica e assim contribuir para as dificuldades na recuperação de informações e acesso aos recursos.

5.1 NECESSIDADE INFORMACIONAL NO DOMÍNIO DA ONCOLOGIA

> Daqui a 50 anos toda a gente vai ter cancro e mais do que um, mas não vamos morrer disso (Sobrinho-Simões, 2016).[30]

A história da Medicina, como a conhecemos atualmente, decorre desde a Antiguidade, com a realização de cirurgias complexas pelos egípcios. Já os gregos avançaram no aspecto do diagnóstico. Estes são apontados como os pioneiros no estudo dos sintomas das doenças. Destaca-se, neste contexto, a figura de Hipócrates, considerado o pai da Medicina.

Durante a Idade Média, devido à atuação da Igreja, não ocorreram grandes avanços na Medicina (“História da Medicina,” n.d.). Nesse período a Medicina era vista como uma prática mágica ou, quando religiosa, associada aos milagres. Era um período que imperava a intolerância religiosa, o obscurantismo e o misticismo, conhecido também como "idade das trevas". Destaca-se nessa altura o papel dos monges

[30] Manuel Sobrinho-Simões, fundador do IPATIMUP – Porto, Portugal, considerado em 2015 o patologista mais influente do mundo, em entrevista para o expresso.pt (http://expresso.sapo.pt/podcasts/a-beleza-das-pequenas-coisas/2017-02-24-sobrinho-simoes-daqui-a-50-anos-toda-a-gente-vai-ter-cancro-e-mais-do-que-um.-mas-nao-vamos-morrer-disso). A afirmação do pesquisador está associada às investigações que promovem diagnósticos precoces e tratamento eficazes.

beneditinos que se preocupavam com a doença e o seu tratamento, período em que teriam surgido os primeiros hospitais. Uma das primeiras escolas de Medicina fundada na Europa, no século X, foi a de Salermo, de forte influência muçulmana, que se destacou como um marco no período cristão pela formação laica dos médicos.

Só a partir do Renascimento – séculos XV e XVI –, houve uma intensa atividade científica que permitiu maior conhecimento sobre o funcionamento do corpo humano. No século XIX, grandes foram os avanços, época em que as bactérias começaram a ser estudadas por Louis Pasteur.

Atualmente, a Medicina conta com inúmeras tecnologias que propiciam qualidade na prevenção, no diagnóstico precoce e no tratamento. Não apenas a Medicina curativa se tornou consistente, como a preventiva tomou uma dimensão que permite uma melhor qualidade de vida para muitas pessoas em todo o mundo ("História da Medicina," n.d.).

O avanço na Medicina, como em outros domínios, é decorrente de investigações e geração de conhecimento. Tem-se, assim, dois aspectos bem delineados em relação a esse domínio: a Medicina como domínio de conhecimento – uma ciência –, e uma Medicina como domínio de aplicação desse conhecimento (profissões médicas). Aspectos que orientaram o desenvolvimento e a utilização de algumas práticas, como a Medicina baseada em evidência, que busca conectar esses dois aspectos: ciência e prática por meio do método científico que envolve técnicas e investigação científica como a melhor via de tratar os pacientes ("Medicina," 2015). A informação passa a ser crucial para um atendimento bem-sucedido, para o ensino, para o desenvolvimento profissional, educação e orientação do paciente.

Por sua vez, a Oncologia é uma especialidade médica que relaciona de forma intensa a ciência com a prática. Ela pode ser definida como

> [...] a especialidade médica que estuda as neoplasias, ou tumores, benignas ou malignas. A origem da palavra vem do grego *onkos*, que significa volume, que sabemos atualmente ser causado pelo acúmulo de células cancerosas em regiões do organismo.
> Os tumores podem ser localizados ou invadir outros tecidos, disseminando-se por todo o corpo. A Oncologia estuda a forma como o câncer se desenvolve no organismo e determina o tratamento mais adequado para cada caso (Centro de Combate ao Câncer, 2014).

Trata-se de uma disciplina complexa que abrange especialidades como Cirurgia, Pediatria, Patologia, Radiologia, Psiquiatria, Física e Biologia. A investigação em relação ao câncer cobre a investigação básica que envolve, entre outras, Bioquímica, Biologia Celular e Genética; e áreas aplicadas, como Psico-

Oncologia e Epidemiologia. Já a investigação clínica relaciona-se com a Oncologia Médica, Oncologia Cirúrgica, Radioterapia, Oncologia Diagnóstica, entre outras (The Netherlands Cancer Institute Antoni van Leeuwenhock Hospital, 2010).

Observam-se avanços no domínio da Oncologia em decorrência de investigações em Medicina nuclear, nas técnicas de Imagiologia e na sequenciação do genoma humano. Esses avanços são decorrentes do conhecimento resultante da investigação multidisciplinar entre físicos, químicos, engenheiros, matemáticos, biólogos, geneticistas e clínicos que se refletem na prática clínica (Costa, 2010).

Os primeiros registros sobre a Oncologia remontam ao antigo Egito, por volta de 3.000 a.C, associados a Imhotep, um sacerdote egípcio. O sacerdote descreveu 45 casos referentes à patologia, embora o conceito mais similar ao empregado atualmente tenha sido utilizado por Hipócrates, em 460-375 a.C. O câncer só passou a ser estudado como patologia a partir do século XIX (Hajdu, 2011). O primeiro tratado médico da Antiguidade sobre câncer é associado a Galeno (131-203 d.C). Esse autor buscou analisar a etiologia da doença (Costa, 2010).

Em Portugal, os primeiros registros sobre o câncer foram publicados entre a Idade Média e o Renascimento – Tratado das Epidemias, em 1418, por Valesco de Taranta, em que o autor relatava experiências de um cirurgião que extirpava tumores. Identifica-se ainda no período renancentista a publicação Centurias (1511-1568), do médico português José Rodrigues de Castelo Branco, grande conhecedor da obra de Galeno, sobre tratamento oncológico em que prevalecia o tratamento paliativo ou a intervenção cirúrgica na fase inicial da doença (Costa, 2010).[31]

Os principais acontecimentos em relação à Oncologia se deram durante o século XVIII, quando foram adotadas técnicas cirúrgicas baseadas em estudos anatômicos mais precisos e os casos de câncer passaram a ter relevância científica com o desenvolvimento de investigações em relação à etiologia, como a identificação de fatores predisponentes, como o tabaco e o rapé (Costa, 2010). Atualmente, é vista como uma especialidade estratégica que trata de questões de alto impacto, como a social e a econômica. Os investimentos necessários para essa área são vultosos e emergentes e requerem alta tecnologia e pessoal qualificado (psicólogos, fisioterapeutas, físicos, geneticistas, biologistas, patologistas, hematologistas, entre outros (American Society of Clinical Oncology, 2016; Direcção-Geral da Saúde. Direcção de Serviços de Planeamento, 2012). Essa vertente multidisciplinar é um desafio a mais, já que os médicos necessitam de conhecimentos técnicos e científicos dos envolvidos para serem

[31] Registros sobre a história da Oncologia em Portugal poderão ser consultados na publicação *Luta contra o cancro e a Oncologia em Portugal*, do autor Rui Manuel Pinto Costa (COSTA, 2010).

capazes de emitir diagnósticos e recomendar tratamentos com o maior grau de segurança possível (Direcção-Geral da Saúde, 2012).

Pode-se dizer que, além da multidisciplinariedade, a quantidade de informações e dados disponíveis é um fator que exige, por parte dos profissionais e investigadores, uma constante atualização e interação entre a investigação e a prática clínica (Direcção-Geral da Saúde, 2012).

O número crescente de pacientes em todo o mundo – EUA, 14,5 milhões de pessoas com câncer; EU, projeção de 2,010.765 (American Society of Clinical Oncology, 2016; De Azambuja, Ameye, Paesmans, Zielinski, Piccart-Gebhart, & Preusser, 2014) –, influi no aumento da busca por informações tanto por profissionais para responderem questões de atendimento ao paciente, quanto por parte dos pacientes que, a cada dia, procuram por mais informações e tornam-se mais questionadores (Fourie & Bakker, 2009).

Por outro lado, temos um aumento exponencial de informações que é marcado por ser fragmentada em termos de tipos, formatos, vocabulários e sistemas de informação, projetando-se como um vasto campo para o desenvolvimento de investigações. Como citado, o domínio da Oncologia engloba vários outros domínios que também são de alta complexidade no que se refere à organização e à recuperação da informação. Como exemplo de um desses domínios, pode-se citar a Biologia Molecular, na qual, dentre os tipos de informações utilizados, incluem-se dados de sequências genômicas e genéticas. Dados sobre as funções e interações de moléculas complexas também são crescentes (MacMullen & Denn, 2005).

Esse domínio ainda possui outras duas características marcantes: a dificuldade em delimitar a abrangência da informação médica e a dispersão do conhecimento em distintas subdisciplinas. Segundo Sobrinho-Simões (2016, p. 6), "[...] a compreensão das bases celulares do cancro e o domínio da linguagem oncológica são indispensáveis para a construção de uma estratégia adequada para a tal luta contra o cancro, tanto em termos individuais, como sociais". Isso porque "[...] os conceitos em Oncologia são muito complexos". Diante dessas dificuldades, novas formas de indexação e classificação de artigos estão sendo exploradas no intuito de garantir a recuperação dessas informações, independentemente do domínio ou da especialidade. Projetos, como classificação automática de assuntos relacionando citações (Pudovkin & Garfield, 2002), uso de anotações semânticas legíveis por máquinas (Dinakarpandian, Lee, Vishwanath, & Lingambhotla, 2006; Marcondes, 2011) são alguns dos exemplos. MacMullen & Denn (2005) defendem que, no atual contexto dos objetos biológicos, como as micropartículas e os genes, é necessário o desenvolvimento de ontologias e a padronização da nomenclatura e de metadados para que possa existir interoperabilidade entre diferentes repositórios e sistemas de informação.

5.1.1 Oncologia: perfil dos profissionais

O domínio da Medicina, como se descreveu, envolve duas categorias distintas de profissionais: aqueles que atuam diretamente com o cuidado da saúde dos pacientes, como médicos; e os investigadores, que têm como objetivo estudar e analisar aspectos, fatores, substâncias e ocorrências relacionados com a saúde das pessoas. Ou seja, essas duas categorias exercem suas funções de modo a convergir na promoção e cuidados com a saúde.

O número de médicos que trabalham com a Oncologia nos EUA, correspondente ao ano de 2015, é de 22.431; na União Europeia (UE), um relatório de 2008 previu que existiriam 9.453 oncologistas em 12 países nesse mesmo período (De Azambuja et al., 2014). Em Portugal, os dados estatísticos disponibilizados pela Ordem dos Médicos de Portugal indicam o total de 193 médicos oncologistas que atuavam na área e, desse total, 163 exerciam funções em instituições vinculadas ao Serviço Nacional de Saúde (2015).

No que se refere a dados estatísticos, a UE possui um diferencial em relação a países como os EUA e o Brasil, a Oncologia Médica só foi reconhecida como especialidade médica a partir de 2011, como parte da diretiva europeia 2005/36/EC sobre a livre circulação de profissionais, o que trouxe, como consequência, um número reduzido de fontes de informação, como relatórios e *surveys* sobre esse domínio. Portugal e França são citados como os países que mais possuem e disponibilizam relatórios, enquanto as informações sobre a Bélgica e o Leste Europeu são praticamente inexistentes (De Azambuja et al., 2014). De Azambuja et al. (2014) identificaram que os oncologistas enfrentam longas horas de trabalho, estão submetidos a forte carga emocional, recebem baixos salários e são altamente pressionados para publicar trabalhos científicos.

Os dados descritos nos estudos supracitados não permitem um aprofundamento no que se refere aos perfis profissionais dos médicos e investigadores, mas alguns autores mostram como o médico contribui para o exercício do investigador e vice-versa. Cochrane (1972) se posiciona em relação à eficácia e à eficiência na prática clínica no Serviço Nacional de Saúde americano e afirma que os resultados de uma investigação podem colaborar para a melhoria da prática clínica. Por outro lado, os investigadores utilizam-se de questões e ocorrências que surgem na prática clínica para o desenvolvimento de investigações científicas (Owen, 1995, Machado, Rebholz-Schuhmann, Freitas, & Couto, 2015).

Em virtude das diferenças existentes entre a prática clínica e a investigação, é observado que as NIs de ambos possuem suas especificidades (Ely, Burch, & Vinson, 1992, McKibbon & Fridsma, 2006). No âmbito da investigação científica, os investigadores trabalham em um contexto amplo de tipos e fontes

de informação. Kim (2010a) exemplifica essa afirmação ao declarar que os investigadores biomédicos não restringem suas buscas aos artigos científicos; elas se estendem às bases de dados, repositórios, bibliotecas, registros hospitalares, imagens clínicas e a informações provenientes de leigos, como indica Chalmers (1995). Já os médicos obtêm e utilizam dados a partir das consultas com os pacientes, acesso aos sistemas de informações de unidades de saúde e prontuários (Gorman, 1993, Prakasan, 2013, Smith, 1996) e interação com seus colegas (Koopmans, 2002, Liverman, Ingalls, Fulco, & Kipen, 1997, Smith, 1996, Timpka, Ekström, & Bjurulf, 1989). No entanto, para alguns autores, há um aumento no uso da literatura científica pelos médicos como forma de eles obterem respostas às questões clínicas, atualização profissional e aquisição de conhecimento e assim proporcionar um atendimento aos pacientes com melhor qualidade (Clarke, Belden, Koopman, Steege, Moore, Canfield, & Kim, 2013, Gorman, Ash, & Wykoff, 1994, Lee, 2013, McKibbon & Fridsma, 2006, Osheroff & Bankowitz, 1993, Wyatt, 2000).

Outro exemplo claro da relação investigação x prática clínica é observada por meio da investigação translacional, que integra a investigação básica com a investigação clínica e possibilita a aplicação prática do conhecimento produzido (Drolet & Lorenzi, 2011). A investigação translacional tem suas origens associadas ao Instituto Nacional de Câncer dos Estados Unidos (NCI) (Guimarães, 2013).

Os investigadores, em parte, utilizam-se de insumos provenientes da prática clínica no desenvolvimento de suas investigações, artigos, *papers* apresentados em eventos científicos e relatos clínicos (Zawawi & Majid, 2001). As contribuições dos relatos clínicos para o ensino e a investigação médica têm sido registradas anteriormente a Hipócrates. Como exemplo, pode-se citar o relato clínico publicado em 1981 na American Journal of Dermatopathology sobre o que é denominado atualmente de Síndrome da Imunodeficiência Adquirida (AIDS/SIDA). É importante ressaltar que a divulgação de relatos clínicos por relevantes publicações científicas repercute no aumento da quantidade e da qualidade desse tipo de comunicação (McCarthy & Reilly, 2000).

Frei (1982) relaciona o progresso da investigação sobre o câncer com a integração da prática e a investigação médica, fato que se ajusta à proposta deste estudo: alimentar essa integração a partir da incorporação dos TNIs, assegurando a semântica por meio de elementos de metadados, de forma a dirimir possíveis problemas em relação à descoberta de informações científicas (Figura 19).

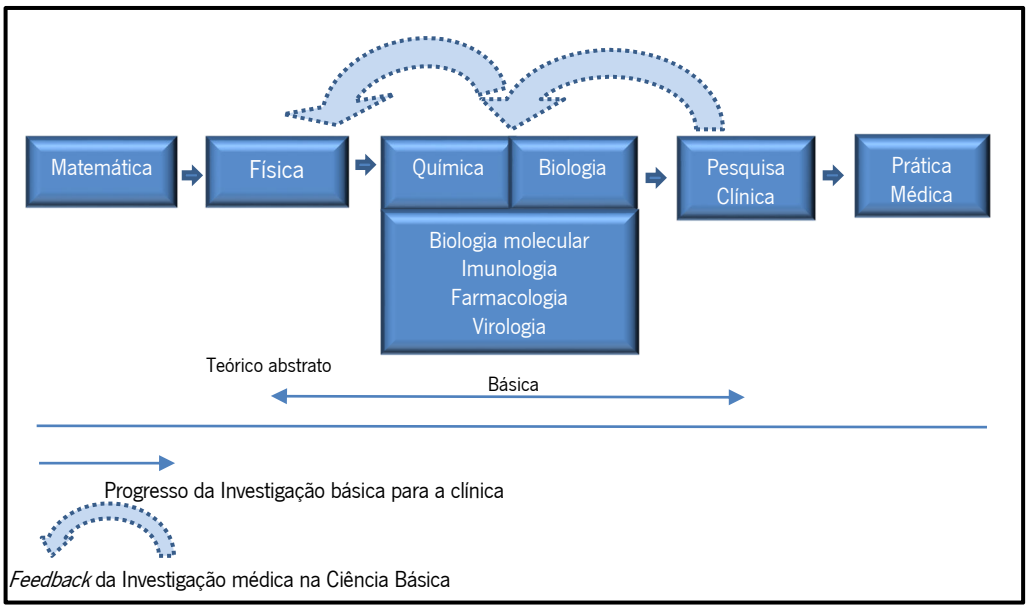


Figura 19. Relacionamento entre a clínica e a investigação básica. Recuperado de Clinical cancer research: an embattled species de E. Frei, 1982, Cancer, 50, p. 1982.

Ainda nesse sentido, observa-se que as duas categorias (médicos e investigadores) integram o Ciclo de Vida da Informação: criação, aquisição, catalogação/identificação, armazenamento, preservação e acesso à informação (Hodge, 2000) no domínio da Oncologia. Tem-se, assim, uma descrição de possíveis atividades desenvolvidas pelos médicos e investigadores associadas ao Ciclo de Vida da Informação:

a) criação: o processo de criação corresponde ao "ato de produzir o produto de informação". O produto de informação, no caso dos médicos, são os prontuários, registros clínicos e publicação de relatos de casos clínicos, por exemplo. Os investigadores, por sua vez, publicam artigos, *papers*, dissertações, teses, patentes, entre outros. Para Hodge (2000), nesta fase o autor contribui na descrição do recurso (inclusão de metadados) já no início do ciclo de vida da informação.

b) aquisição: a atividade de aquisição e coleta é o estágio em que o objeto criado é "incorporado" fisicamente ou virtualmente ao arquivo". Nessa fase, os médicos e investigadores podem participar ativamente com o depósito de publicações nos RDCs.

c) catalogação/identificação: "Uma vez que o objeto digital é incorporado à coleção, é necessário identificá-lo e catalogá-lo." A identificação digital proporciona o endereço para localização e possibilita a ligação com outros recursos. A catalogação com o uso de metadados possibilita a organização, o acesso e a preservação dos recursos. Nessa fase os médicos e investigadores podem contribuir com descrições de acordo com as especificidades do seu domínio.

d) armazenamento: o armazenamento dos recursos em formato digital traz inúmeras vantagens e, no caso do domínio da Medicina em particular, a rapidez no armazenamento e na publicação e também a diversidade de tipos de recursos (imagens, texto, dados) contribuem para os avanços no diagnóstico, tratamento e prevenção.

e) preservação e acesso: estabelece como os recursos são gerenciados em termos de preservação e forma. Assim, esses profissionais podem ter acesso às informações por longo tempo.

A produção e a disseminação de informação fazem parte de um contexto que aponta para a necessidade de ações que visam a integrar esses profissionais no que tange à organização, gestão e recuperação dos recursos informacionais.

> *Biomedical research and the practice of medicine*, *separately and together*, *are reaching an inflection point: the capacity for description and for collecting Data*, *is expanding dramatically*, *but the efficiency of compiling*, *organizing*, *manipulating these Data — and extracting true understanding of fundamental biological processes*, *and insights into human health and disease*, *from them— has not kept pace. There are isolated examples of progress: research in certain diseases using genomics*, *proteomics*, *metabolomics*, *systems analyses*, *and other modern tools has begun to yield tangible medical advances*, *while some insightful clinical observations have spurred new hypotheses and laboratory efforts. In general*, *however*, *there is a growing shortfall: without better integration of information both within and between research and medicine*, *an increasing wealth of information is left unused* (Council, 2011).

Nota-se que esses dois tipos de profissionais, embora possam ter TNIs distintos, podem convergir para a identificação dessas necessidades com o objetivo de beneficiar o exercício das duas práticas profissionais – médico e investigador –, a partir da busca e da recuperação de informações nesse domínio.

5.1.2 Características das necessidades informacionais de médicos e investigadores no domínio da Oncologia

A busca pela identificação das NIs dos médicos teve significância reconhecida a partir de 1972 (Forsythe, Buchanan, Osheroff, & Miller, 1992) (não foi encontrada menção semelhante em relação aos estudos relativos aos investigadores do domínio das Ciências da Saúde). Na literatura, observa-se que há um número maior de investigações voltadas à identificação das NIs dos médicos, em grande parte, desenvolvidas no final da década de 80 e início dos anos 90 (Forsythe et al., 1992, Gorman, 1995, Gorman et al., 1994, Haynes, Wilczynski, McKibbon, Walker, & Sinclair, 1994, Osheroff & Bankowitz, 1993, Osheroff, Forsythe, Buchanan, Bankowitz, Blumenfeld, & Miller, 1991, Smith, 1996, Timpka,

Ekström, & Bjurulf, 1989, Woolf & Benson, 1989). A partir de 2000, houve um aumento no número de investigações, inclusive com a preocupação de utilizar o conhecimento sobre as NIs dos profissionais de saúde para o desenvolvimento de sistemas de informação hospitalares (Allen et al., 2003, Collins, Currie, Bakken, & Cimino, 2009, Ely, Osheroff, Gorman, Ebell, Chambliss, Pifer, & Stavri., 2000a, González-González, Escortell Mayor, Hernández Fernández, Sánchez Mateos, Sanz Cuesta, & Riesgo Fuertes, 2007, Jerome, Giuse, Gish, Sathe, & Dietrich, 2001, Prakasan, 2013) e até mesmo para viabilizar a integração desses sistemas de informação com as BDBs (MacMullen, 2003).

Este estudo trabalha com a perspectiva de Taylor (1962), que enfatiza a relação entre a tecnologia e o utilizador, desde que esse último tenha consciência da sua NI. Assim, partiu-se do princípio de que o médico e o investigador têm consciência das NIs que subsidiam o desenvolvimento das suas atividades, e a tecnologia deve promover meios para que essas necessidades sejam atendidas. No entanto, apesar do contínuo avanço da tecnologia, esta ainda depende de fatores que proporcionam diferentes resultados em relação à qualidade da resposta que o utilizador obtém. Um desses fatores é a disponibilização de ferramentas e de informações que atendam à demanda dos seus utilizadores. Para isso é preciso identificar que tipo de necessidade de informação os médicos/investigadores buscam ao consultarem os serviços de informação, a *Web* ou um artigo científico?

Observa-se que há escassez de estudos com ênfase nos TNIs dos investigadores (Andrade & Baptista, 2014), principalmente no domínio da Medicina. Quando se restringe à especialidade Oncologia, esse problema se torna ainda mais evidente (Fourie & Bakker, 2009). Pelo fato de a Oncologia ser um domínio do tipo "guarda-chuva", em que outras disciplinas estão envolvidas, e pelo número reduzido de investigações nesse contexto, aqui são analisados trabalhos que abordam outras especialidades médicas. Estudos referentes às áreas mais recentes de investigação, como Biologia Molecular e Genética, são identificados em maior quantidade, ainda assim, sob um número modesto. Dos estudos existentes, muitos exploram as NIs em relação às fontes bibliográficas (artigos, livros) e aos dados (Grefsheim & Rankin, 2007). Percebe-se também que esses trabalhos não se restringem apenas ao contexto das bibliotecas, dos repositórios digitais e das bases de dados bibliográficas, pois já existem estudos cuja preocupação está voltada para os repositórios de dados (Kim & Gilbertson, 2007).

Embora não seja objetivo desta investigação analisar aspectos referentes a esse tipo de repositório, considera-se que a organização dos dados por vezes se manifesta por meio de **figuras** dentro dos artigos científicos, em que os dados são tratados assumindo forma de fluxogramas, gráficos de barra, diagramas,

sequências de **genes** e **árvores filogenéticas**[32] (White, 2010), itens que, no contexto deste estudo, estão associados aos TNIs. Para Kim & Gilbertson (2007), na Oncologia, a complexidade na obtenção das informações pelos investigadores está relacionada em parte com a diversidade dos tipos de dados, o que, por si só, já é um fator que deveria suscitar a investigação dos TNIs dos investigadores.

Segundo MacMullen & Denn (2005), questões pertinentes à área de atuação dos biologistas são por vezes muito específicas. Os autores esclarecem que poucos são os estudos que buscam conhecer o comportamento e identificar as necessidades e o uso de informação nesse domínio.

No ambiente clínico, as dificuldades relativas à tradução de questões dos profissionais para os sistemas médicos ou bases de dados bibliográficas, repositórios, adicionada a restrição de tempo, dificultam não apenas a atualização e aquisição de conhecimentos, mas também comprometem o acesso a informações que podem subsidiar o atendimento a um paciente (Jerome et al., 2001, Maggio, Steinberg, Moorhead, O'Brien, & Willinsky, 2013, Osheroff & Bankowitz, 1993). Essas dificuldades têm merecido esforços de alguns estudiosos no que se refere à identificação dos TNIs desses profissionais (Del Fiol, Workman, & Gorman, 2014, Liverman et al., 1997). A realidade atual se mostra tímida em relação a estudos sobre a identificação dos TNIs da comunidade científica, apesar de a literatura mostrar que os recursos *on-line* e os recursos eletrônicos, quando disponíveis, e os recursos tradicionais de informação dão aos médicos acesso ao conhecimento, que possibilita a melhoria do atendimento do paciente e a qualidade no tratamento (McKibbon & Fridsma, 2006).

Há alguns anos, os médicos não se restringem apenas à consulta de periódicos (Gorman et al., 1994); eles buscam informações disponibilizadas em bases de dados bibliográficas, como a PubMed e a UpToDate, para fornecer respostas clínicas aos seus pacientes (Maggio et al., 2013). Esses médicos se fundamentam na Medicina Baseada em Evidências (EBM) para emitir diagnósticos e indicar tratamentos a seus pacientes (Agoritsas, Merglen, Courvoisier, Combescure, Garin, Perrier, & Perneger, 2012, Maggio et al., 2013). De acordo com Lin & Fushman (2005), os médicos, aos buscarem por EBM, podem utilizar quatro elementos para auxiliar nas pesquisas por artigos científicos: problema, intervenção (**tratamento**), comparação e resultado (PICO). Ou seja, questões clínicas ou de investigação podem ser formuladas utilizando esses quatro componentes. Questões sobre problemas podem ser exemplificadas da seguinte forma: qual a causa do problema? Corresponde à **etiologia** e a **fatores de risco**. Qual a frequência do problema? Frequência. Essa pessoa tem um problema? **Diagnóstico**. Quem pode ter

[32] Árvore filogenética é uma representação gráfica, em forma de árvore, apresentando as relações evolutivas entre várias espécies ou outras entidades que possam ter um ancestral comum (Wikepedia, 2018).

esse problema? **Prognóstico** (http://learntech.physiol.ox.ac.uk/cochrane_tutorial/cochlibd0e84.php). Essas questões são exemplos de TNIs.

Ao analisar a literatura, foi possível identificar algumas diferenças dos TNIs de médicos e investigadores no domínio da Medicina. Os médicos possuem características específicas em relação aos TNIs para determinados atendimentos clínicos. Quando se trata de urgências clínicas, por exemplo, a demanda é por informações facilmente acessíveis, informações resumidas sobre um determinado **produto químico** e seus efeitos e informações detalhadas sobre o tipo de tratamento a ser adotado. Já os que exercem atividades relacionadas com o desenvolvimento de políticas necessitam de informações que respaldem a tomada de decisão, o que requer um aprofundamento da busca para obtenção de fontes que forneçam sustentação às suas decisões (Liverman et al., 1997). Nos anos 90, alguns autores apontavam indícios de que os sistemas de informação não conseguiam penetrar na prática rotineira dos médicos por estarem associados à ausência de estudo das NIs para sua concepção (Smith, 1996). Atualmente, já existem estudos que apontam para um crescimento no uso de **evidências clínicas**, por exemplo, para implantação de políticas públicas (Brasil, 2015); por outro lado, estudos demonstram que ainda há dificuldade no acesso a trabalhos que auxiliem à tomada de decisão (Campbell, Redman, Jorm, Cooke, Zwi, & Rychetnik, 2009, Menon & Stafinski, 2005).

Mais uma vez, isso demonstra a relevância em investigações e ações no sentido de entender os TNIs dessa comunidade a fim de proporcionar resultados de busca mais relevantes. Afinal, se Glasziou et al., citados por Agoritsas et al. (2012), estiverem corretos, ao afirmarem que "[...] o uso de motores de busca agora é tão essencial como o estetoscópio", pode-se imaginar a relevância que os sistemas de informação estruturados como as bases de dados bibliográficas e os repositórios digitais ocupam nesse contexto.

É sob essa perspectiva que se propõe a realização deste estudo, ou seja, o desenvolvimento de um PAM baseado em TNIs de médicos e investigadores no domínio da Oncologia para aplicação no contexto dos RDCs, como forma de otimizar a interoperabilidade semântica entre os RDCs e outros sistemas (LOD), o que, consequentemente, proporciona maior qualidade ao processo de busca e recuperação da informação.

Para identificação dos TNIs, foram usadas duas estratégias, a análise da literatura e o estudo de caso em que foi desenvolvido um *Survey*, com utilização de questionário eletrônico anônimo.

Os procedimentos metodológicos adotados para o desenvolvimento do estudo incluíram a revisão sistemática da literatura[33]. A pesquisa bibliográfica utilizou palavras-chave pré-definidas para realizer buscas nos *sites* de pesquisa e com motores de busca, com o objetivo de identificar os TNIs de médicos e dos investigadores que atuam no domínio da Medicina/Oncologia. A opção por utilizar os termos *information needs*, *medicine* e *health science*, *oncology* ou *cancer* foi baseada em estudo anterior desenvolvido pelas autoras Andrade & Baptista (2014), em que se verificou a escassez de investigações sobre NIs no domínio da Medicina, principalmente em Oncologia. As bases de dados selecionadas para consulta foram a Scopus, Medline, *Web* of Science (WOS), Library, Information Science and Technology Abstract (LISTA) e Library and Information Science Abstracts (LISA), justificadas pela relevância como bases multidisciplinares, no domínio das Ciências Saúde e na Ciência da Informação, respectivamente, e o motor de busca Google Scholar. A pesquisa realizada no Google Scholar considerou apenas os 200 resultados mais relevantes. A Figura 20 mostra o processo de pesquisa bibliográfica adotado.

O processo de busca proporcionou a recuperação de 3.969 itens (Tabela 54). Após análise do título e *abstract* com base no critério de inclusão e exclusão, esse número foi reduzido para 251 artigos. Artigos que necessitavam de leitura na íntegra para identificação da sua relevância não foram considerados.

O critério de inclusão e exclusão foi estabelecido em vista do objetivo do estudo. Foram excluídos artigos que: a) descrevessem sistemas de informação e/ou tecnologias sem que fossem mencionados aspectos em relação às NIs dos utilizadores; b) focassem exclusivamente os técnicos da área da saúde (enfermagem/atendente de consultórios/técnicos de saúde); c) analisassem o comportamento informacional; e d) abordassem NIs informacionais referentes às fontes de informação e ao uso.

A partir desses 251 artigos, foram recuperados mais 110, provenientes de citações ou artigos que os referenciaram, o que totalizou 361 itens. Os 361 artigos foram reanalisados, seguindo ainda o mesmo procedimento (análise do título e *abstract*), em que se obteve um subtotal de 99 artigos para leitura na íntegra. Seguindo o critério de inclusão e exclusão, pertinência e relevância para a consecução do objetivo proposto (identificar TNIs de médicos/investigadores que atuam no domínio da Medicina e/ou Oncologia), foi possível selecionar 23 artigos para análise qualitativa por meio do *software* NVivo. Esse número foi alterado para 28 após atualização do processo de busca em julho de 2015, quando também foi adicionado novo termo, *information requirements*.

[33] Revisão sistemática "[...] é a investigação científica que reúne estudos relevantes sobre uma questão formulada, utilizando o banco de dados da literatura que trata sobre aquela questão como fonte e métodos de identificação, seleção e análises sistemáticos, com intuito de se realizar uma revisão critica e abrangente da literatura" (Wikipédia, 2018).

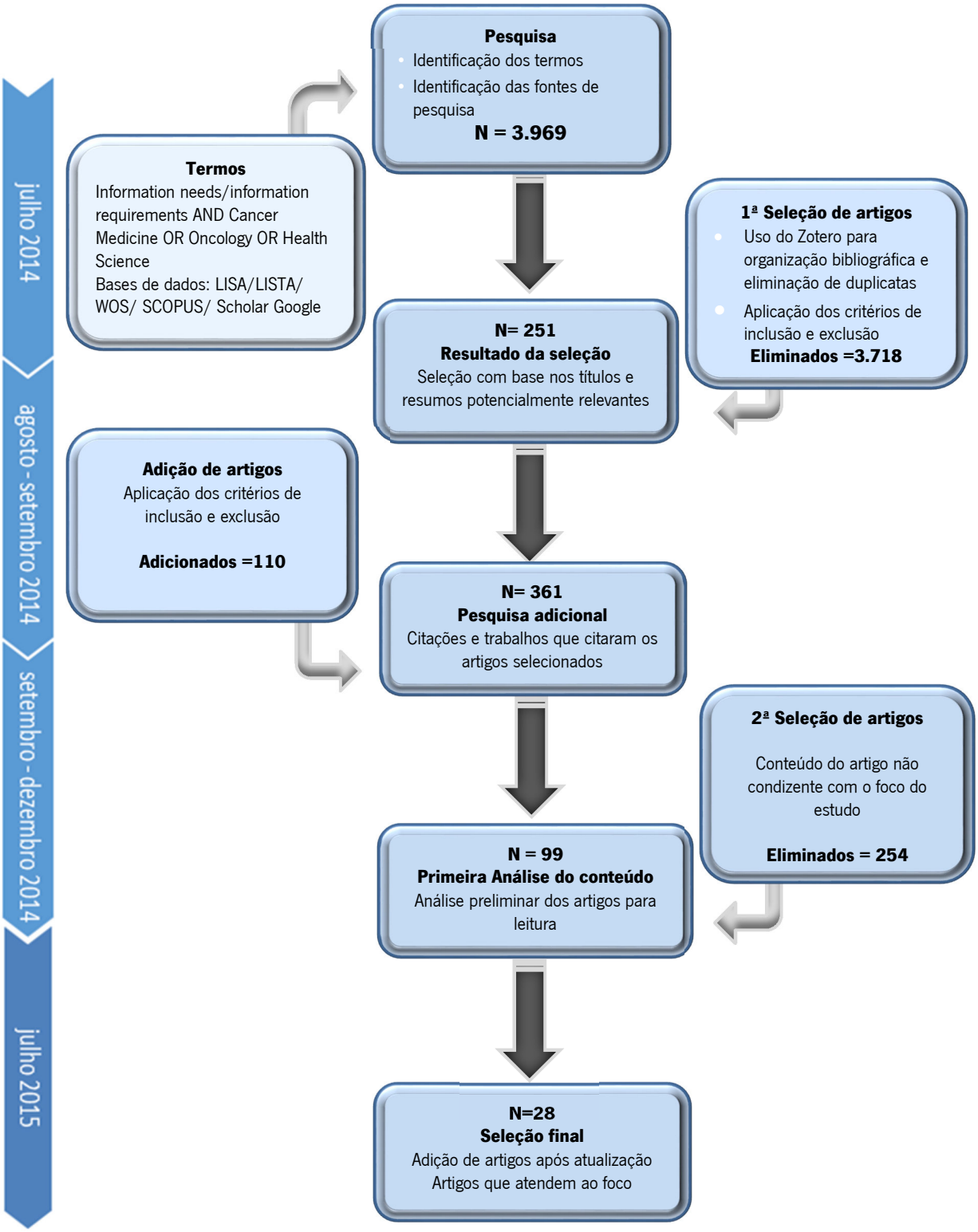


Figura 20. Representação gráfica do processo de pesquisa bibliográfica.

Tabela 54. **Necessidades informacionais: artigos identificados**

Base de dados	Campo	Termos	Refinamento	Artigos recuperados	Artigos Selecionados
Web of Science	Tópicos	Information needs and medicine or health	• Information science library science or medical informatics or social sciences interdisciplinary or education educational research or social sciences biomedical or dentistry oral surgery medicine or education scientific disciplines • Artigos	633	39
Scopus	Keywords	Information needs and medicine or health	• Artigos e revisão de literatura	380	8
Scopus	Keywords	Information needs and oncology or cancer	• Artigos e revisão de literatura	108	1
Scopus	Keywords	Information needs and physician or researchers	• Artigos e revisão de literatura	23	9
LISA	Subject	Information needs and medicine or health	• Artigos	23	11
LISTA	Subject	Information needs and medicine or health	• Revistas acadêmicas	683	41
PUBMED*	MESH	Access information and research	• Artigos • Human	659	23
PUBMED*	Título	Information needs and clinicians or researchers or physician	• Artigos • Human • Journal article • Review • Systematic review	650	59
PUBMED*	Título	Information needs and oncology or cancer	• Artigos, Journal article • Review • Systematic review	210	3
Scholar Google	Título e abstract	Information needs and oncology or cancer	• Not patients or patient or parents • Artigos	200 itens mais relevantes	6
Scholar Google	Título e abstract	Information needs and medicine or health and article and researcher or clinician	• Not patients or patient or parents • Artigos	200 itens mais relevantes	18
Scholar Google	Livre	Information needs and health science and researchers or physician	• Not patients or patient or parents • Artigos	200 itens mais relevantes	33
Total 3.969					251*

Nota: *Após eliminação de títulos duplicados.

A relevância em identificar os TNIs é apontada por vários artigos em diversos períodos da história, visto que essa identificação pode proporcionar mudanças no atual paradigma da organização, descrição e recuperação da informação. Tais mudanças podem ser viabilizadas, como relatado, por inclusão de elementos de metadados (Hjørland, Nielsen & Williams, 2001, Jerome et al., 2001) que reflitam as necessidade dos seus utilizadores, não se restringindo à inclusão de elementos de metadados preestabelecidos por padrões que não consideram essas necessidades (Bishop, 1998, Hoffman, 2009, Weinberg, 1988).

Parte dos TNIs dos médicos e investigadores apresentada em alguns trabalhos foi baseada em questões submetidas aos serviços de informação (Allen et al., 2003, Collins et al., 2009, Divoli, Wooldridge, & Hearst, 2010, Gorman, Ash & Wykoff, 1994, Kim & Gilbertson, 2007, MacMullen & Denn, 2005) ou obtidas por meio de instrumentos de coleta de dados, como questionários e entrevistas (Gorman, 1995, Gorman, Ash & Wykoff, 1994, Kim, 2010b, Kim & Gilbertson, 2007, MacMullen & Denn, 2005, Osheroff & Bankowitz, 1993, Timpka, Ekström, & Bjurulf., 1989, Woolf & Benson, 1989).

Este estudo alinha-se à posição de vários autores ao afirmarem que a NI dos utilizadores é um tópico que suscita análise em relação ao contexto e aos domínios (Julien & Duggan, 2000, Markey, 2007a, 2007b, Rowlands, 2007). Foco deste estudo, a Oncologia, assim como os domínios da Biomedicina e da Biologia Molecular, traz consigo desafios devido à diversidade de tópicos, às fontes de informação, aos tipos de dados e à fragmentação das informações geradas pelo próprio domínio e pela multidisciplinaridade (MacMullen & Denn, 2005, Kim, 2010b, Kim & Gilbertson, 2007, Liverman et al., 1997).

Esta fase do trabalho teve por objetivo responder à seguinte questão: que TNIs os médicos ou investigadores possuem em relação a um artigo científico ao consultar uma base de dados bibliográfica ou um repositório digital? Esse é um questionamento que, diante da literatura consultada, é apontado como pouco explorado, não só em relação aos médicos (Gorman, 1995, Woolf & Benson, 1989), mas, sobretudo, com referência aos investigadores (Korjonen-Close, 2005).

O uso do NVivo proporcionou a análise qualitativa dos textos, nomeadamente em relação às seções resultados e conclusão, sobre TNIs a partir da definição e utilização de nós (categorias). Primeiramente, recorreu-se a uma das suas ferramentas, o módulo de consulta por frequência de palavras, para estabelecer possíveis nós. Essa consulta indicou o uso das 100 palavras mais frequentes, eliminando as palavras fora do contexto, adjetivos e verbos (Figura 21).



Figura 21. Nuvem de palavras com os termos mais frequentes nos artigos selecionados.

A partir dessa análise qualitativa, foi possível perceber que alguns termos são comuns aos dois objetos de investigação (***diagnostic procedure***, ***disease***, ***epidemiology***, ***family health history***, ***incidence***, ***patient care***, ***patient information***, ***prognosis***, ***therapy***, ***treatment***). No caso dos TNIs relacionados com o domínio da Biologia Molecular, nota-se que os investigadores procuram informações mais granulares, diferentemente dos médicos que buscam informações mais gerais. Observa-se que os investigadores formulam questões mais específicas, como no caso de tópicos como **genoma** (Kim, 2010), corroborando o pensamento de Bishop (1998), Leen(2013), Sandusky & Tenopir (2008).

Evidencia-se, por meio dessa análise, a concentração de questões no que se refere ao diagnóstico e à terapêutica. Entretanto, quando se reporta aos componentes de publicação (***gen bank***, ***table***, ***caption image***, ***citation analysis***, ***figures***) (Kim & Gilbertson, 2007), observa-se que alguns tipos vão além dos comumente identificados nos repositórios digitais. Em relação aos formatos multimídia, como imagens, a solicitação de especificidades também é revelada (***illumination***, ***acquisition*** *and* ***processing***, ***acquisition context***, ***image center point coordinates sequence***, ***processing***; ***type***, ***file location equipment***, ***3D structure of a protein***) (Kim & Gilbertson, 2007, MacMullen & Denn, 2005).

A literatura mostra ainda que alguns autores salientam a importância de identificar dados como **Gene ID** (Ácido desoxirribonucleico-DNA) ***tumor size***, ***statewide large cancer registry***, ***outcomes and biomarkers***, em relação a algumas doenças, como o câncer (Kim & Gilbertson, 2007). Kim (2010b)

defende que a Biomedicina engloba a interdisciplinariedade e que investigações sobre genomas envolvem dados além dos textuais e necessitam de informações sobre **amostras biológicas** (tipos de tecido e sangue). Dessa forma, nota-se que a diversidade de TNIs apresentada vai ao encontro das críticas formuladas por Bishop (1998), Weinberg (1988) e Hjørland, Nielsen, & Williams (2001), em relação ao processo de indexação e ao estabelecimento de pontos de acesso, que são padronizados.

Esses termos são alguns exemplos de refinamento de informação requerido por investigadores. Já os médicos, ao acessarem os serviços de informação, não formulam questões com esse grau de requisito, até mesmo pelas dificuldades encontradas em traduzir suas questões para os sistemas (Osheroff & Bankowitz, 1993).

Quando foram analisados os TNIs de médicos e investigadores, por meio da opção busca de palavras mais frequentes do NVivo, puderam ser visualizados alguns dos tipos mais adotados, assim como sua influência nas atividades dos dois profissionais. As Figuras 22 e 23 apresentam nuvens de palavras que ilustram alguns TNIs comuns aos médicos e investigadores, assim como tipos associados apenas aos médicos ou aos investigadores. Foram detectados os seguintes tipos comuns aos dois profissionais: ***treatment***, ***finding***, ***patient***, ***test***, ***management***, ***drug***, ***epidemiology***, ***diagnostic***, ***data***, ***disease***, ***incidence***, ***interpretation*** **e** ***prognosis***. Os tipos associados apenas aos médicos são: ***cause***, ***etiology***, ***effects adverse***, ***orientation***, ***etiology***, ***cost***, ***exam***, ***education*** e ***biopsy***. Já aos investigadores: ***demographic***, ***expression***, ***protein***, ***health***, ***history***, ***method***, ***species***, ***specimen***, ***tissue*** *e* ***sample.***



Figura 22. Nuvem de palavras com TNIs dos médicos.



Figura 23 Nuvem de palavras com TNIs dos investigadores.

Os artigos utilizados na análise qualitativa para identificação dos TNIs são apresentados na Tabela 55. A primeira coluna elenca os autores analisados; a segunda informa os domínios; e a terceira traz os objetos participantes.

Tabela 55. **Artigos selecionados para identificação dos TNIs**

Artigo	Domínio	Objetos da investigação
Allen et al. (2003)	Clinical information system	Clinicians
Bjerre, Paterson, McGowan, Hogg, Campbell, Viner, e Archibald (2013)	Continuing medical education	Physicians
Bowden, Kromer, & Tobia (1994)	Medical library (MEDLINE)	Physicians
Bryant (2000)	Library practices	Family doctors
Cheng (2004)	Clinical questions	Healthcare professionals
Collins et al. (2009)	Clinician information system	Clinicians, medical residents, nurse practitioners, nurses, pnysician assistant
Dee & Blazek (1993)	Hospital library	Rural physicians
Del Fiol, Workman e Gorman (2014)	Patient care	Clinicians
Divoli, Wooldridge e Hearst (2010)	Search engine	Bioscientists
Dogan,et al. (2009)	PubMed	PubMed
Ebell, Cervero, & Joaquin (2011)	Point of care	Health care professional
Ely et al. (2000a)	Knowledge bases	Family physicians
Ely, Osheroff, Ebell, Bergus, Levy, Chambliss, & Evans (2000b)	Primary care	Doctors
Giuse, Huber, Giuse, Brown, Bankowitz, & Hunt (1994)	AIDS (HIV)	Health care professionals in HIV
González-González, Dawes, Sánchez-Mateos, Riesgo-Fuertes, Escortell-Mayor, Sanz-Cuesta, & Hernández-Fernández (2007)	Primary care in Madrid, Spain	Primary care physicians
Gorman (1995)	Clinical information system	Physician
Gorman, Ash & Wykoff (1994)	Bibliographic database	Primary care physicians
Graber, Randles, Monahan, Ely, Jennissen, & Anderson (2007)	Emergency unit	Physicians
Haynes et al. (1994)	Information resources	
Jerome et al. (2001)	Scholarly paper	Medical librarians
Lin & Fushman (2005)	Information Systems	Clinicians, physicians
Kim & Gilbertson (2007)	Cancer center researchers	Cancer researchers, leading researchers, biomedical researchers, leading researchers
Kim (2010b)	Bio-repository information systems	Medical doctors, Pathologist, Tissue bankers
Osheroff & Bankowitz (1993)	Internal medicine practice	Physician
Prakasan (2013)	Scholarly paper	Medical faculty members, general practtioners, nursing professinals, family physicians, clinicians
Smith (1996)	Scholarly paper	Doctors
Timpka, Ekström, & Bjurulf (1989)	Medical information	General practioners, primary health care
Woolf & Benson (1989)	Academic medical center	Physicians, interns, residents

A análise desses artigos proporcionou a identificação dos TNIs médicos e investigadodres associados às especialidades Medicina da Família, Cuidados Primários, Patologia, Biociências e Genética (Tabela 56 e 57).

Tabela 56. **Tipos de NIs de investigadores no domínio da Medicina.**

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Divoli, Wooldridge & Hearst (2010)	Biologist	This article focuses on another way to improve bioscience literature search: provide the user with the ability to search over full text, including figure captions, and display the associated figures alongside search results.	Usability study	—	Abstract Analysis technique Author name Biological process Biological process analysis technique Biological process and species tissue type Captions Cell compartment Cell compartment drug Cell compartment species name Chemical Common dna sequence Equipment Experimental results Figures Gene Method Method species name Organic compound species name Plasmid experimental protein Protein name Protein structure Protein subunit Species disease Species name Species tissue type Table Title

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Dogan et al. (2009)	PubMed users'	Investigate of user interactions observable through one month of PubMed logs.	Log analysis		Gene/Protein Disorder Chemical/Drug Biological Process Medical Procedure Living Being Research Procedure Cell Component Body Part Tissue Device
Kim (2010)	Biomedical scientists and researchers, Junior researchers	This study aimed to investigate the characteristics of biomedical scientists and their information requirements, including biological materials as well as data requirements.	Questionnaire	Descriptive analysis, cross-tabulation, Two Step Cluster analysis, SPSS	Amount of tissue Amounts of biosample Anatomic *sites* Anatomic *sites* of interest Biosample requesting criteria Biosample storage mediums Cost willing to spent Demographic Demographic condition Demographic/general caracteristics patients Demographic/socioeconomic Demographics Disease names Disease of interest Diseases Family health history Frequency of biosample requests General health conditions Genetic & proteomic data (disease biomarker expression, protein expression) Health behavioural data
Kim (2010)					

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Health outcome Histopathoogy findings Laboratory findings Laboratory test interpretation Large scale genomic data (tma, cdma microarray) Needs of associated data Needs of biosamples Physiology findings Primary or metastatistic Private health information Processing & storage Progression and recurrence Radiologic findings Required Sample annotation Sample collection Sample distribution Sample processing methods Sample-specific elements Specimen Specimen preparation and preservation method Tissue sources Total number of specimens needed Treatment and response Types of bio-samples Types of biosample processing
Kim & Gilbertson (2007)	Cancer researchers, leading	Investigate the information requirements of biomedical researchers at a large	Interviews, query analysis, analysis of publications	—	Absolute gene expression (signals) (gene) Administration Amount of antigen present (gene)

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Kim & Gilbertson (2007)	researchers, biomedical researchers	and complex cancer center through interviews, query analysis, analysis of publications, and conceptual data modeling.			Amount of target detected (gene) Anatomic frame of reference (image) Anatomic location of target detected (gene) Anatomic region Antibody (gene) Antigen (the protein) (gene) Biological function (gene) Cancer directed/non directed (therapy) Caption (image) Cell intensities (gene) Chemical agent administration Chip design (gene) Chip name (gene) Citation analysis Class of disease Clinical path Coding regions (cds) Color coding (image) Color depth (image) Comparative gene expression (comparison, statistic) (gene) Comparison (tumor v adjacent normal) (gene) Data on large populations Date of procedure Demographic Description of processing (image) Diagnostic molecular Diagnostic procedures Disease Disorder-based query Epidemiology Equipment (image)

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Kim & Gilbertson (2007)					Expression analysis results (gene) Expression level (gene) Extent of disease Family clinical history File location equipment (image) File name (image) Filne size (image) Functional test Funding resource Gene expression Gene expression experiments (gene) Gene id Gene id (DNA) Genetic information Histologic type Histology Homology Illumination (image) Image Image acquisition and processing (image) Image acquisition context Image acquistion module Image center point coordinates sequence Image class Image instance Image processing Image size Image type Incidence Influence of occupational IRB approval

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Kim & Gilbertson (2007)					Lab findings Lab results Lab test values Laboratory data Locus Magnification (image) Manufacturer (image) Method of confirmation Method of diagnosis Methodology (gene) Methodology (affymetrix, northern blotting, tacman) (gene) methodology (gene) Molecular Molecular observation Morphologic expression signatures in tissue Morphological (image) Morphological query Mortality Name of equipment (image) Name of gene Name of protein Nature of disease No of rows & cols of cells (chip design - gene) Nuclear atypia Organ primary Outcomes and biomarkers Outcomes/progression Pathologic diagnosis Pathologic diagnosis (histological types) Pathologic findings Pathology Pathology specimen diagnosis
Kim & Gilbertson (2007)					

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Kim & Gilbertson (2007)					Patient care Patient clinical history Patient consent Patient demographic Patient diagnosis Patient information Patient outcomes People performed Personal lifestyle factors on cancer incidence and survival Physical description (image) Pixel resolution (image) Population-based questions Primary investigator Primary magnification (image) Probe id (probe set) Probe type (gene) Prognosis Progression Project Protein expression Protein expression experiment Protein id Protein target Protocol parameter (gene) Public health issue basedPublication (author, title, journal, title, volume, number, year, pages, pubmed id, gen bank)

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Race and gender Ratio (image) Related image sequence information Research Results of gene Results of proteomic studies Sample type Sequence Size of pixel (image) Slide coordinate module *Software* (image) Specific grading method Specimen Specimen acquisition and processing Personal lifestyle factors on cancer incidence and survival Statewide large cancer registry Structure Survival data Teaching/didactic Test value units Therapy Therapy history Tissue microarray (gene) Tissue processing Tissue sample Title (image) Title of project Transcript (probe set) (gene) Transcript id (rna) Treatment

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Tumor size Type of specimen (block, frozen, tissue, slides, serum) Types of procedure Types of test Type of physiscian Visual findings Weight of specimen Where the antigen is present (gene)

Tabela 57. **Tipos de NIs de médicos no domínio da Medicina.**

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Allen et al. (2003)	Clinicians	To determining what information needs arise, in a given context, based on factors such as user traits, patient traits, CIS task being performed, and specific data that may be triggering the need.	*Survey* instrument, analyses of system log files	—	Anion gap Disease Laboratory test Medication Procedure Prognostic Radiology reports Report Report finding Sign Symptom Test Test result
Bjerre et al. (2013)	Primary care practitioners	To identify areas of perceived greatest needs for continuing medical education (CME)	Questions using Web-based platform or handheld device	Quantitative content analysis methods	Could finding y be caused by drug x? How should I manage condition/finding/situation y? How should I treat finding/condition y (given situation z)? Is drug x (or drug class x) indicated in situation y or for condition y? Is drug x safe to use in situation y? Is test x indicated in situation y? (includes screening tests) Is x a risk factor for condition y? Should this kind of patient get prophylactic drug x to prevent condition y? What is the cause of test finding x? What is the usual course (or natural history) of condition y?

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Bowden, Kromer, & Tobia (1994)	Physician	To examine the efficacy of a circuit librarian program designed to improve the availability and use of medical information in four counties of the lower Rio Grande Valley.	Questionnaire		Diagnosis Drug information Lab tests Physical signs/symptoms Provide info to family Referral Treatment
(Bryant, 2000)	Family doctor	To explore the information needs and information seeking behaviour of family doctors, identifying any differences in attitudes and behaviours deriving from membership of a training practice and investigating the impact of a practice librarian.	Interviews and group discussions; analysis of all recorded use of the library		Diagnosis Pharmacological information Prognosis Therapeutic
Collins et al. (2009)	Clinicians, medical residents, nurse practitioners, nurses, pnysician assistant	To examine, in a controlled setting, the usefulness of streamlining information retrieval at the point of care.	Morae *software* (audio e visual capture)	—	Can drug x cause (adverse) finding y?* How is this done at this institution? How should i manage condition x (not specifying diagnostic or therapeutic)? How should i treat condition How should i treat condition x (not limited to drug treatment)? Test indication What are my patient data? What is normal in this situation? What is the cause of test finding x? What is the dose of drug x What is the drug of choice for condition x? What is the patient education related to x?

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Collins et al. (2009)					What is this (unfamiliar) domain-specific information? What test is indicated in situation x?
Cheng (2004)	Clinicians	This article presents the results of complementary research studies on the behaviors of hospital clinicians in asking clinical questions and the relationship between asking of questions, outcome of information searches, and success in problem solving	Interviews Questionnaire Randomized Controlled Study	Categorization according to the components (patient or problem, intervention, outcomes)	Diagnosis information Drug therapy Equipment Interventions (therapy) Prognosis Technology Therapy Equipment Technology
Dee & Blazek (1993)	Rural physicians	To describe the information needs and the information-seeking behavior of rural physicians	Interviews and observation Review of patient charts	–	Diagnosis Etiology Psychological Treatment
(Del Fiol et al., 2014)	Clinicians	To systematically review studies that examined the questions clinicians raise in the context of patient care decision making	Systematic review		What is the drug of choice for condition X? What is the cause of symptom X? How should I treat condition X (not limited to drug treatment)? What is the cause of physical finding X? What test is indicated in situation X? What is the dose of drug X? Can drug X cause (adverse) finding Y? What is the cause of test finding X? Could this patient have condition X? How should I manage condition X (not specifying diagnostic or therapeutic)? What is the prognosis of condition X? What are the manifestations of condition X?

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					What conditions or risk factors are associated with condition Y?
Ebell, Cervero, & Joaquim (2011)	Physicians	To identify the clinical questions that health care professionals have at the point of care and explore whether these questions could be used to drive a needs assessment for continuing education programs.	Observation Questions submitted by *e-mail*, text message, web site for mobile devices	The questions were stored in a Microsoft Excel spreadsheet and were assigned a category from the Generic Questions Taxonomy and a clinical category (cardiovascular, psychiatric) by the principal investigator (M.H.E).	Cause of clinical finding Prevention Prognosis Test selection Treatment selection
Ely et al. (2000a)	Family doctors, primary care doctors, family doctors	To develop a taxonomy of doctors' questions about patient care that could be used to help answer such questions. Our goal was to produce a logical and concise classification scheme that could be applied reproducibly to the full range of questions that occur in primary care.	Recorded questions between patients and doctors	-	Administration Administration in face of adverse effects (drug prescribing) (treatment) Adverse effects (drug prescribing) (treatment) Availability (drug prescribing) (treatment) Body part (anatomy) on physical exam or imaging study (diagnosis) Causation/ association (etiology) (epidemiology) Cause/ interpretation of clinical finding (diagnosis) Community services (management) Condition (name finding) Condition/ finding (management)

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Ely et al. (2000a)					Continuing medical education (provider) (education) Cost (drug prescribing) (treatment) Cost (diagnosis) Course/ prognosis (epidemiology) Criteria/ manifestations (diagnosis) Dosage (drug prescribing) (treatment) Drug prescribing (treatment) Efficacy/ indications/ drug of choice (drug prescribing) (treatment) Ethics Etiology (epidemiology) Findings caused by drug/adverse effects of drug (drug prescribing) (treatment) Genetics (epidemiology) How to advise (doctor-patient communication) (management) How to approach difficult issue (doctor-patient communication) (management) How to do it (treatment) How to prescribe (treatment) Inconsistencies (diagnosis) Indications/ efficacy (test) (diagnosis) Interactions (drug prescribing) (treatment) Legal Mechanism of action (drug prescribing) (treatment) Name finding (drug prescribing) (treatment) Name finding (test) (diagnosis) Orientation (test) (diagnosis)Orientation/ composition (drug prescribing) (treatment)

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Ely et al. (2000a)					Patient (education) Patient compliance (doctor-patient communication) management) Pharmaco-dynamics/ absorption (drug prescribing) (treatment) Physical characteristics (drug prescribing) Prevalence/ incidence (epidemiology) Prevention (drug prescribing) (treatment) Prevention (efficacy/indications) (treatment) Referral (management) Risk factors/ disease agents (epidemiology) Safety/contraindications (includes pregnancy and breast feeding) (adverse effects) (drug prescribing) Serum levels (drug prescribing) (treatment) Sign (cause/interpretation of clinical finding) Symptom (cause/interpretation of clinical finding) (diagnosis) Test (lab, ecg, imaging, biopsy, skin test, element of physical exam, etc) (diagnosis) Test (lab, ecg, method imaging, biopsy, skin test, element of physical exam, etc) (diagnosis) Test (lab, ecg, preparation imaging, biopsy, skin test, element of physical exam, etc) (diagnosis) Test (lab, ecg, timing/imaging, biopsy, monitoring, skin test, element of physical exam, etc) (diagnosis) Test (lab, skin test, biopsy, imaging, element of physical exam, etc) Timing (treatment) Trainee (education) (nonclinical)

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Treatment (efficacy/indications) (treatment)
Ely et al. (2000b)		To characterize the information needs of family physicians by collecting the questions they asked about patient care during consultations and to classify these in ways that would be useful to developers of knowledge bases	Interview	The questions were categorized as either topics or generic questions	What is the cause of symptom X? What is the dose of drug X? How should I manage disease or finding X?\ Could this patient have disease or condition X? How should I treat finding or disease X? Is drug X indicated in situation Y? Is test X indicated in situation Y? What is the cause of physical finding X? What is the cause of test finding X? What is the drug of choice for condition X?
Giuse et al. (1994)	Health care professionals	To examine the information needs of health care professionals in HIV related clinical encounters, and to determine the suitability of existing information sources to address those needs.	Patient charts	Categorized by MeSH	Adverse effects of drug therapy Clinical trial protocol Description of disease Diagnosis/etiology Diagnostic procedure Diet therapy Disease complication Drug information Efficacy of diagnostic procedure Efficacy of drug therapy False-negative rate of diagnostic procedure Morbidity/mortality of diagnostic procedure Patient counseling Patient information/education Patient workup Treatment protocols/regimens

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
González-González et al. (2007)	Primary care physicians	To determine the information needs and information-seeking patterns of a random representative sample of primary care physi- cians seeing patients in practices where consultations are of short duration.	Questionário	SPSS v. 11.0	Condition/finding Cost Course Drug prescribing Epidemiology Etiology Incidence Inconsistencies Interpretation of clinical fi nding Management Name finding Orientation Physician-patient communication Prevalence Prognosis Test (laboratory, ecg, imaging, biopsy, skin test, element of physical examination, etc) Type of study
Gorman (1995)	Physicians –	This article reviews research on the information needs of physicians to examine methodological problems, summarizes what is known, and suggests directions for future research.	Interview	—	Logistic information Medical knowledge Patient data Population statistics Social influence
Gorman, Ash, & Wykoff (1994)	Primary care physicians	To answer the question, "If primary care physicians used the journal literature to pursue all of their clinical questions, for how many of these questions could	Interviews searches of bibliographic data-bases clinician feedback	—	Anonymous provides, less expense than inpatient treatment? anti-inflammatory effect Are nonacetylated salicylates really safer-and how much safer-in patients with NSAID GI

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Gorman, Ash, & Wykoff (1994)		they find an answer, and what would be the cost of finding and providing this information?			intolerance (who benefit from anti-inflammatory effect)? At what age is screening prostate-specific antigen [testing] indicated in a low-risk patient? At what point is endoscopy indicated in patients with esophagitis who remain symptomatic on medication? Diagnosis Exams Exams effects Family history For diagnosis of deep-vein thrombosis, how good is ultrasound; does it obviate the need for venogram (can it exclude the diagnosis)? Is amoxicillin safe for use in a lactating woman? Is hypothyroidism associated with high cholesterol or low? Low-risk patient Patient education Risk Testing What is [sic] the sensitivity and specificity of arterial ultrasound exam of the lower extremities? What is the dose of Imferon? What is the exact increase in risk of thrombotic events on oral contraceptives in a woman with family history of myocardial infarction (her grandmother at age forty-nine) and of deep-vein thrombosis?

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Where can I send this patient for education about his alcoholism: more education than Alcoholics
Graber et al. (2007)	Physicians	To categorise questions that emergency department physicians have during patient encounters.	Observational study Recorded verbatim	Categorized using a taxonomy of clinical question (Ely et al. 2000)	Administrative rules Aetiology, NOS agents Diagnosis, characteristics of a test Diagnosis, how is test A done? Diagnosis, laboratory interpretation Diagnosis, multiple symptoms Diagnosis, name finding Diagnosis, sign interpretation Diagnosis, symptom Diagnosis, symptoms of a disease Diagnosis, using laboratory clinically Diagnosis, what is syndrome Q? Diagnosis, what is the name of that body part? Epidemiology, aetiology Epidemiology, aetiology: risk factors and disease Epidemiology, course and prognosis Epidemiology, prevalence Ethics Interpretation Management Treatment NOS treatment X risk versus benefit Treatment, adverse reactions, safety and contraindications Treatment, drug composition Treatment, drug cost

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
Graber et al. (2007)					Treatment, drug identification Treatment, drug interactions Treatment, drug of choice, indications, efficacy Treatment, drug physical characteristics Treatment, drug, availability Treatment, drug, pharmacokinetics Treatment, drug, serum levels Treatment, how do I treat condition X, how good is Treatment, how do you administer drug X? Treatment, how do you do procedure X? Treatment, how long to treat? Treatment, prophylaxis Treatment, side effects Treatment, what is the dose of drug X?
Haynes et al. (1994)	Medline	To develop optimal MEDLINE search strategies for retrieving sound clinical studies of the etiology, prognosis, diagnosis, prevention, or treatment of disorders in adult general medicine	Search terms and phrase in Medline	—	Diagnosis Etiology Prognosis Treatment
Jerome et al. (2001)	Medical librarians	To examine the types of questions received by Clinical Informatics Consult Service (CICS) librarians from clinicians on rounds and to analyze the number of clearly differentiated viewpoints provided in response.	Analysis of questions on CICS Knowledge database	Subject analysis	Diagnosis Diagnostic procedure Differential diagnosis Disease complication Disease description Disease prognosis Education Etiology Patient information

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Treatment Treatment adverse effect Treatment efficacy
Lin & Fushman (2005)	Clinicians, physicians	This paper presents a case study on how existing resources and knowledge in the clinical domain can be integrated into a rich computational model of the information seeking process.	Case study	—	Comparison Diagnostic methods selection Differential diagnosis Intervention Outcome Problem
Prakasan (2013)	Medical faculty members, general practtioners, nursing,professinals, family physicians, clinicians	To review the scholarly papers published on the topic information needs and use of healthcare professionals and related fields	Review the scholarly	—	Diagnosis Drug therapy Ethical issues Patient care Personal safety
Osheroff & Bankowitz, (1993)	Medical information	To gather descriptive data about the use of medical information software by physicians who were early users of medical resources.	Questionnaire	Categorization	Articles on a specific topic Disease information (general information about a disease Hx, PE, and/or lab for a given disease Drug information Drug side effects Finding of differential diagnosis General information about a disease How a patient should be managed disease information Information about test or clinical finding Literature citations A specific article Articles on a specific topic Natural history of a disease

Autores	Objeto da Investigação	Objetivo	Método de Coleta De Dados	Método de Análise de Dados	Tipo de Necessidade de Informação
					Relationship among clinical states Utility of diagnostic test When or how should drug or therapy be used drug-drug interactions
Smith (1996)	Doctors	This paper examines studies of the needs and wants of doctors for medical knowledge; I pay less attention to the other information needs and wants of doctors.	Systematic review	—	Substance to treatment
Timpka Ekström & Bjurulf (1989)	General practitioners	To investigate the GP's decision-making habits with respect to his need and search for information, his present use for sources of information, and a description of his "information dilemma", his perceived need for information.	Questionnaires Critical incidents	SPSSX using multiple regression and chi-square analyses	Cause Condition finding Diagnosis Epidemiology Etiology Incidence Interpretation of clinical finding Name finding Prevalence Prognosis Therapy Therapy control
Woolf & Benson (1989)	Physicians	To describe in greater detail the information-seeking behavior and information needs of physicians, and to compare the information requirements of physicians at different levels of training.	*Survey*, questionnaire, interview	Statistical and graphical analysis (DEC Vax 11/780	Basic science Diagnostic criteria Differential diagnosis Drug information Lab test interpretation Physical signs interpretation Symptom interpretation Treatment modalitie Treatment recommendation

Os TNIs apresentados nas Tabelas 56 e 57 foram analisados e avaliados de forma a: a) reduzir a ambiguidade, redundância e imprecisão; e b) classificá-los em categorias. A partir das definições de termos presentes no vocabulário MESH e em dicionários especializados, foi possível atender aos dois objetivos. Em adição, utilizaram-se a Taxonomia de Ely et al. (2000b) e a árvore hierárquica do MESH como parâmetro para categorização dos TNIs (Tabela 58).

Tabela 58. **Tipos de necessidades de informação de médicos e investigadores que atuam no domínio da Medicina**

Categoria	Tipos Necessidades Informacionais	Categoria	Tipos Necessidades Informacionais
Epidemiologia	Causation/ association (etiology) Course/ prognosis (epidemiology) Etiology (epidemiologia) Genetics (epidemiology) Population statistics (epideiologia) Prevalence/ incidence (epidemiology) Risk factors/ disease agents (epidemiology)	**Assistência à saúde**	Clinical trial information Clinically condition (orientation) Community services (management) Condition/ finding (management Doctor-patient communication (management) How to advise (management) How to approach difficult issue (doctor-patient communication) (management) Management (not specifying diagnostic or therapeutic) (management) Patient compliance (doctor-patient communication) (management) Practices of other providers (management) Specifying diagnostic or therapeutic
Patologia	Cause and risk factors Class of disease (cancer) (nature of disease) Data of diagnosis (nature of disease) Date of staging (extent of disease) Diagnosis/etiology Disease complication Disease prognosis Documentation for this cancer) (nature of disease) Method of confirmation (clinical Method of diagnosis (nature of disease) Organ primary (class of disease) Pathologic diagnosis (histological types) Precursor lesions (nature of disease) Specific findings tnm (stage) (extent of disease) Stage (extent of disease) Staging system (stage) (extent of disease) Status at staging (extent of disease) Tnm stage (extent of disease)	**Pacientes: monitoramento**	Autopsy Cancer status at death Date Date last known alive Date last verified recurrence–free Date of death Date of first recurrence Date of last contact Date of metastasis Demographic data elements Ethics Family history of disease How a patient should be managed How documented (biopsy, physical exam) Metastases (multiple) Method used to determine first recurrence Place of death Primary cause of death Race and gender Quality of survival Site of first distant recurrence Site of local–regional recurrence Site of metastasis Type of first recurrence (local, distant) Vital status

Categoria	Tipos Necessidades Informacionais	Categoria	Tipos Necessidades Informacionais
Diagnóstico	Anatomic region (specimen) (diagnostic procedure) Body part (anatomy) on physical exam or imaging study Cause/ interpretation of clinical finding Characteristics of a test Cost Criteria/ manifestations Date of procedure Diagnosis, multiple Diagnosis, using laboratory Diagnostic aid Disease accuracy (test) Diagnostic criteria Diagnostic procedures Differential diagnosis How processed (specimen) (diagnostic procedure) Image (specimen) (diagnostic procedures) Images (diagnostic procedure) Inconsistencies (diagnosis) Indications/ efficacy (test) Laboratory findings (images) Laboratory and physiology findings Method (test) Morbidity/mortality of diagnostic procedure Name finding (test) Orientation (diagnosis) Pathology findings (specimen) (diagnostic procedure) Patient diagnosis Physical signs interpretation Preparation (test) Sign (cause/interpretation of clinical finding) Specimen (diagnostic procedures) Symptom (cause/interpretation of clinical finding) (diagnosis) Test (lab, ecg, imaging, biopsy, skin test, element of physical exam, etc) (diagnosis) Test (orientation) Test finding (lab, ecg, imaging, biopsy, skin test, etc.) (diagnosis) Timing/ monitoring (test) (diagnosis) Tumor grade, size and stage at dx and resection Type of specimen (block, frozen, tissue, slides, serum) (diagnostic procedure) Types of procedure (radiology, lab, surgery) (diagnostic procedure) Unspecified findings or multiple categories of findings (diagnosis)	**Genética**	Amount of target detected Analysis technique and species Anatomic location of target detected Biological function (protein expression) Biological process and species (genetic) Cell compartment and protein structure Cell component Chemical/organic compound Comparison (tumor v adjacent normal) Expression level (protein expression) Gene expression experiments Gene id (dna) Gene id (genetic information) Gene/protein name and species and cell compartment Methodology (affymetrix, northern blotting, tacman) (genetic information) (probe array & a single stranded dna or rna sequence) Methodology (genetic information) Methodology (protein expression) Name of protein (protein expression) Plasmid Probe type (genetic information) Processing e storage Protein expression (genetic information) Protein expression experiment (genetic information) Protein id (genetic information) Protein target (genetic information) Sample annotation Sample collection Sample type (protein expression) (genetic information) Sample-specific elements Sequence (gene id (dna) Sequence (protein id) Species and cell compartment Species name and author Species name and method Structure (gene id (dna) Structure (protein id) Tissue microarray (anatomic location of target detected (genetic information) Tissue type and disease Transcript (probe set) (genetic information) Transcript id (rna)

Categoria	Tipos Necessidades Informacionais	Categoria	Tipos Necessidades Informacionais
Imagiologia	Anatomic frame of reference (image acquisition module) Chemical agent administration (image acquisition module) Color coding (images) Color depth (images) Compression (images) Description of processing (image processing) Equipment (images) File location (physical description) (images) File name (physical description) (images) File size (physical description) (images) Illumination (image acquisition module) Image acquisition context (image acquisition module) Image acquisition module Image center point coordinates sequence (slide coordinate module) (images) Image type (images) (light microscope, gross, radiography, endoscopy, surgery, electron microscope) Lossy/lossless (images) Magnification (image acquisition module) Manufacturer of equipment (images) Monochrome2 or RGB or YBR_full_422 (images) Name of equipment (images) Physical description (images) Pixel resolution (the amount of specimen subtended by a pixel) (images) Primary magnification (images) Ratio (images) Related image sequence information (images) Size of pixel (images) Slide coordinate module (images) *Software* (equipment) (image) *Software* (image processing) Specimen acquisition and processing (image acquisition module)	**Terapêutica**	Administration in face of adverse effects Adverse effects Availability (treatment) Cancer directed/non cancer directed (therapy) Chemical/drug Contraindications Cost (drug prescribing) Diet therapy Dosage (treatment) Drug prescribing (treatment) Efficacy/ indications/ drug of choice False-negative rate of diagnostic procedure Findings caused by drug/ adverse effects of drug How to do it (treatment) Interactions (drug prescribing) Mechanism of action (drug prescribing) (treatment) Modification,additions/omission to therapy (therapy) Name finding (drug prescribing) Orientation/ composition (drug prescribing) (treatment) Outcomes and progression data – vital status and progression markers Pharmaco-dynamics/ absorption (drug prescribing) Physical characteristics (drug prescribing) Prevention (treatment) Principles/ rationale (treatment) Prior metastatic/recurrent therapy (therapy) Prior therapy (therapy) Protocol (therapeutic procedures) Reactions, safety and Reason for no-therapy Reason for therapy ended Safety/ contraindications (includes pregnancy and breast feeding) Serum levels (drug prescribing) Start/end date (therapeutic procedures) Therapy sequence (therapy) Timing (drug prescribing) Treatment efficacy Treatment recommendation Treatment, adverse Type of therapy (surgery, radiation)

Categoria	Tipos Necessidades Informacionais
Informações sobre os recursos	Título do artigo Título da publicação Autor Assunto Abreviaturas/sigla Figuras Títulos de figura Tipo de estudo Suporte financeiro Factor de impacto Citações bibliográficas Tabelas Títulos de tabelas Notas Análise de citações Ano de publicação Gene

Como comentado anteriormente, encontrou-se um número reduzido de estudos sobre os TNIs de médicos ou de investigadores em diferentes especialidades da Medicina, sendo quase total a ausência desses tipos de estudo no domínio da Oncologia, o que motivou o desenvolvimento de uma investigação nesse domínio.

5.2 *SURVEY* PARA IDENTIFICAÇÃO DOS TIPOS DE NECESSIDADES DE INFORMAÇÃO DE MÉDICOS E INVESTIGADORES NO DOMÍNIO DA ONCOLOGIA

O *Survey* foi desenvolvido a partir de TNIs identificados em estudos anteriores. Após as pesquisas bibliográfica e documental, em que se buscou identificar TNIs no domínio da Oncologia, foi encontrado apenas um artigo que versa exclusivamente sobre esse domínio, mais especificamente em relação à Genética Oncológica. Na ausência de investigações envolvendo a Oncologia, a autora utilizou TNIs identificados por estudos em outras especialidades médicas (Saúde da Família, Cuidados Primários, Medicina Interna e Patologia) para o desenvolvimento do questionário (Tabela 58).

Os TNIs contemplados por artigos clássicos e/ou que apresentassem certo grau de citações (Allen et al., 2003, Divoli, Wooldridge, & Hearst, 2010, Ely et al., 2000a, 2000b, Gorman, Ash & Wykoff, 1994, Jerome et al., 2001, Kim & Gilbertson, 2007, Kim, 2010b, Lin & Fushman, 2005, MacMullen & Denn, 2005, Timpka, Ekström, & Bjurulf, 1989, Woolf & Benson, 1989) foram utilizados como variáveis no questionário. Os 12 artigos foram selecionados com base no diferencial que cada um apresentava em relação aos TNIs.

Quanto aos artigos que traziam TNIs identificados em mais de um estudo ou replicação de estudos anteriores, priorizou-se a utilização dos mais citados no Google Scholar.

A estratégia de investigação adotada para a identificação dos TNIs foi o estudo de caso (Avison & Myers, 2005, Myers, 2013, Yin, 2013), que contempla a identificação de TNIs de médicos e investigadores que atuam na Oncologia em Portugal. Esta investigação tem uma perspectiva predominantemente qualitativa. A análise quantitativa será utilizada para determinar os TNIs mais indicados pelos respondentes que, posteriormente, podem viabilizar a construção de serviços de informação mais eficientes e eficazes no domínio da Oncologia, o que vem ao encontro da afirmação de alguns autores, como Hjørland, Nielsen, & Williams (2001), que na medida que as informações são mais acessíveis elas passam a ser as mais utilizadas.

5.2.1 **A amostra**

Em Portugal Continental, no ano de 2015, a população de médicos oncologistas era constituída por 193 profissionais, dos quais 163 exerciam atividades em instituições do Sistema Nacional de Saúde-SNS (Serviço Nacional de Saúde, 2015). No contexto da Oncologia, o SNS possui as Redes Nacionais de Especialidades Hospitalares e de Referenciação de Oncologia (RNEHR).

As RNEHRs são divididas em região Norte, Centro e Sul e consideram quatro tipos de unidades de assistência em Oncologia (Serviço Nacional de Saúde, 2015):

a) Unidade I – integra os Serviços de Especialidades e Clínicas Multidisciplinares por patologia. Essas instituições oferecem atendimento multidisciplinar que cobre as principais modalidades de tratamento, prevenção, rastreio, diagnóstico, reabilitação e o cuidado terminal. Possuem oncologista permanente, desenvolvem atividades referentes à prestação de cuidados médicos, educação e investigação e participam de grupos internacionais.
b) Unidade II – integra os Centros Clínicos de Oncologia. Essa unidade corresponde ao Departamento de Oncologia dentro de um hospital. Os centros devem prover tratamento oncológico localmente, que inclui cirurgia, radioterapia e quimioterapia, e realizar atividades de formação e de investigação clínica na área.
c) Unidade III – corresponde às unidades de Oncologia que atendem às patologias mais frequentes; e
d) Unidades de proximidade – são afiliadas a um centro da Rede e dão assistência complementar aos pacientes.

A amostra foi não aleatória. Considerou-se a proporção dos recursos humanos na estrutura proposta para as RNEHRs de Oncologia do SNS de Portugal (Serviço Nacional de Saúde, 2015), o papel dos institutos de investigação no contexto nacional e o percentual de retorno dos questionários respondidos. Baseada nesse critério, a Região Norte de Portugal foi selecionada como área geográfica para o estudo, em consideração à concentração de médicos e à relevância dos centros de investigação e hospitais no domínio da Oncologia, no cenário nacional e internacional.

De acordo com o Serviço Nacional de Saúde (2015), os médicos estão distribuídos da seguinte forma: 3, Alentejo; 4, Algarve; 23, Centro; 64, Lisboa e Vale do Tejo; e 69 no Norte. As Unidades na Região Norte estão distribuídas da seguinte forma: uma, Tipo I (IPO, com 27 médicos); duas, Tipo II (Centro Hospitalar de São João, 9 e Centro Hospitalar de Braga, 5); seis, Tipo III (Centro Hospitalar do Porto, 5; Centro Hospitalar de Vila Nova de Gaia e Espinho, 4; Centro Hospitalar de Trás-os-Montes e Alto Douro, 4, Unidade Local de Saúde de Matosinhos, 4; Centro Hospitalar de Entre Douro e Volga, 6; Centro Hospitalar do Algarve, 3); e quatro Unidades de proximidade (Centro Hospitalar de São Tomé, 1; Centro Hospitalar do Médio Ave, 1).

Em relação aos investigadores no domínio da Oncologia da Região Norte, não foi encontrado nenhum registro oficial (https://www.ine.pt/, http://www.dgeec.mec.pt/np4/46/, http://www.pordata.pt), a não ser o número disponibilizado pelos institutos de investigação por meio de *sites* ou por informação verbal, cujo resultado foi: 120 no Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP); 109 investigadores no Instituto de Biologia Celular e Molecular (IBMC); 57 vinculados ao Departamento de Patologia e Oncologia da Universidade do Porto; Instituto Abel Salazar da Universidade do Porto, 48. Há evidências de que investigadores e médicos podem exercer funções em mais de uma instituição. Assim sendo, o total apresentado não corresponde a um número real de respondentes.

A decisão em selecionar apenas uma área geográfica para investigação está relacionada com o fato de se tratar de estudo de caso da população de médicos e investigadores que atuam em Oncologia, cujo recorte geográfico (região norte de Portugal) não determina necessariamente um viés no resultado (Cull, O'Connor, Sharp, & Tang, 2005). A ampliação da investigação para outras regiões não garante um número maior de respostas. Em geral, o índice de resposta obtido por qualquer tipo de *survey* é de 10 a 35% (Falconer & Hodgett, 1999). Considerando que no Norte de Portugal se destacam em nível nacional e internacional quatro instituições no âmbito da Oncologia, dois hospitais (Serviço Nacional de Saúde, 2015) e dois institutos de investigação, a escolha por essa área geográfica torna-se significativa.

As quatro instituições em destaque são:

a) Instituto Português de Oncologia (IPO) – foi criado em 1923, sob denominação de Instituto Português para o Estudo do Cancro em Lisboa. Em 1930, recebeu a atual denominação. Atualmente possui sede em Coimbra e no Porto, IPO-Coimbra e IPO-Porto, respectivamente, e todos passaram a integrar a rede hospitalar a partir de 2002. O IPO-Porto tem se distinguido na investigação científica nas áreas de investigação de transferência e ensaios clínicos. Atualmente é a instituição referência em ensaios clínicos em Portugal.

b) Hospital São João no Porto – é um hospital de referência no tratamento do câncer para os distritos do Norte de Portugal;

c) Instituto de Patologia e Imunologia Molecular da Universidade do Porto (IPATIMUP) – é uma associação privada, sem fins lucrativos e de utilidade pública, vinculada à Universidade do Porto. Foi criado em 1989, possui articulação funcional com o IBMC e o Instituto de Engenharia Biomédica (INEB) e protocolos de colaboração com várias instituições internacionais.

   É referência na investigação de translação e pós-graduação em Oncobiologia, Medicina Molecular e Genética, em Portugal, além de prestar atendimento aos pacientes oncológicos. Na instituição, trabalham investigadores das Faculdades de Medicina, Ciências, Farmácia e do Instituto de Ciências Biomédicas de Abel Salazar da Universidade do Porto (https://sigarra.up.pt/up/pt/*Web*_base.geralpagina?p_pagina=p%C3%A1gina%20est%C3%A1tica%2 0gen%C3%A9rica%20186).

e) Instituto de Biologia Molecular e Celular (IBMC) – é uma instituição de investigação multidisciplinar, criada em 1997, com ênfase em três áreas de investigação: Infecção e Imunidade, Biologia Molecular e Celular e Neurociência. Os recursos humanos reúnem investigadores e docentes da Universidade do Porto e de hospitais universitários que integram 43 grupos de investigação (http://www.cienciaviva.pt/divulgacao/arteeciencia/ibmc.asp).

Ao final de 2015, após início da coleta de dados, o IPATIMUP e o IBMC, juntamente com o INEB passaram a fazer parte do Instituto de Investigação e Inovação em Saúde (I3S), chefiado pela Universidade do Porto, consagrando a colaboração existente entre as quatro instituições. Além dessas instituições, seis escolas da Universidade do Porto e três hospitais (CH São João, Centro Hospitalar do Porto e Instituto Português de Oncologia – IPO) também passaram a contribuir para as atividades desenvolvidas pelo consórcio, organizadas em três unidades temáticas: câncer, resposta do hospedeiro e interação e neurobiologia; e doenças neurológicas. O programa integrativo do consórcio, que visa a fomentar a investigação transdisciplinar, envolve cerca de 800 investigadores de vários domínios do conhecimento (http://www.i3s.up.pt).

A população investigada é restrita à Região Norte de Portugal, que inclui hospitais e institutos de investigação conforme descrito (Tabela 59). A população compreende 69 médicos e 334 investigadores que atuam em hospitais e em institutos de investigação, respectivamente. Salienta-se que existem profissionais que atuam simultaneamente em hospitais e institutos.

Tabela 59. **População a ser investigada: Região Norte de Portugal**

Profissionais	**Hospitais**	**Institutos de investigação**
Médicos da especialidade	69	334
Investigadores	–	
Total	69	**334***

Nota. *Inclui professores, médicos e investigadores do IPATIMUP e IBMC. Recuperado de Serviço Nacional de Saúde de Serviço Nacional de Saúde, 2015, http://www2.portaldasaude.pt/portal/conteudos/a+saude+em+portugal/publicacoes/estudos/gtrelatorio+oncologiamedica.htm, https:www.ibmc.up.pt/ e informação verbal de gestores.

5.2.2 **A coleta de dados**

O instrumento utilizado para coleta de dados foi o questionário eletrônico (via *SurveyMonkey*). A coleta de dados, desde o seu início – julho de 2015 – mostrou-se limitada em relação ao acesso às instituições. O envio de solicitações para associações, instituições de ensino e investigação e hospitais não se mostrou eficaz. Foram enviados *e-mails* para institutos, associações, hospitais e profissionais solicitando espaço para divulgação da investigação, os quais não foram respondidos (APÊNDICE B). Por meio de contato interno, foi possível a apresentação da investigação para uma investigadora do IPATIMUP, que promoveu a divulgação do questionário. O processo de coleta de dados decorreu no período de julho de 2015 a maio de 2016.

Inicialmente, a investigação utilizou a modalidade de disseminação de *link* para o questionário distribuído entre os integrantes do IPATIMUP. Posteriormente, investigadores do próprio instituto enviaram para integrantes do IPO e do Hospital São João. É importante salientar que a utilização de *links* restringe o controle sob o número de pessoas que tiveram acesso ao *link*, assim como a sua identificação, caso o respondente não se identifique em campo específico.

Outros procedimentos também foram adotados para obtenção de potenciais contatos para participação do *survey* :

a) identificação dos profissionais em faculdades de Medicina de universidades portuguesas (Departamento de Patologia e Oncologia da Universidade do Porto; institutos de investigação -Instituto

Abel Salazar da Universidade do Porto; profissionais dos hospitais no Norte (São João, Braga) (total de 123 *e-mails* enviados); e institutos de investigação (IBMC e IPATIMUP, disseminação de *links*);

b) identificação dos profissionais registrados na Plataforma DeGois (Sistema de *Curriculum Vitae* português), que possibilitou o envio de quatro *e-mails*;

c) identificação das bibliotecárias da área da saúde, com o objetivo de intermediar o contato com utilizadores dos serviços de informação onde eles exerciam suas funções para disseminação do *link* para o questionário; e

d) busca na Medline por artigos para identificação de *e-mails* de autores vinculados a instituições portuguesas, que resultou na identificação de 194 *e-mails* (Davies, 2008).

Os *e-mails* identificados de profissionais em Portugal (321) foram analisados a fim de eliminar as duplicações, dos quais 20 foram eliminados, em vista de alguns investigadores integrarem mais de uma instituição. As atividades supracitadas resultaram na Tabela 60.

Tabela 60. **Demonstrativo de respostas em Portugal**

Portugal		
Questionários enviados	Questionários – *e-mail* –respondidos	Questionários – *links*– respondidos
301	28	31

Nota-se que o número de respostas foi aquém da quantidade de questionários enviados. O baixo índice de respostas obtidas por meio de *survey*s eletrônicos pode estar associado ao perfil dos entrevistados, ao domínio a que estão associados e à ausência de incentivos. O índice de respostas com a utilização do *survey* eletrônico varia normalmente entre 10-15% (Lang, 2002, Park & Tosaka, 2010, Manfreda, Bosnjak, Berzelak, Haas, Vehovar, & Berzelak, 2008).

Na Medicina, estudos demonstram um índice que oscila de 10 a 30% de respostas (Aitken, Power, & Dwyer, 2008, Fan & Yan, 2010, Fischbacher, Chappel, Edwards, & Summerton, 2000, VanGeest, Johnson, & Welch, 2007). Trata-se de um domínio que está fortemente relacionado com o uso de incentivos para participação, mas que nem sempre faz com que se ultrapasse os 35% (Aitken, Power, e Dwyer, 2008; Kellerman & Herold, 2001).

Infelizmente, a investigação obteve um baixo índice de resposta, possivelmente pela falta de incentivos, contanto apenas com a boa vontade e a disponibilidade dos respondentes (Fourie & Bakker, 2009).

A autora tentou incluir o Brasil na investigação, entretanto isso não foi possível. O número de respostas do Brasil foi muito baixo em relação à população de profissionais que atuam no país (3.419 em 2015) (Demografia Médica no Brasil, 2015), mesmo diante das seguintes iniciativas:

a) colaboração individual de professores/investigadores para divulgação do *link* para o questionário (Instituto Nacional do Câncer – INCA/RJ, Hospital Santa Rita – Vitória/ES, Sírio Libanês/SP);

b) utilização do *Curriculum Lattes* – Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – para identificação de investigadores, professores e médicos que atuam no domínio da Oncologia para envio dos questionários (203 *e-mails* enviados), já que esse sistema é utilizado de forma mandatória pela maior parte das universidades brasileiras públicas e algumas privadas;

c) envio para institutos de investigação e hospitais (Hospital AC Camargo – 61 *e-mails* enviados; solicitação aos hospitais Albert Einstein, Sírio Libanês); e

c) busca na Medline para identificação de artigos sobre Oncologia nos últimos dez anos, cujos autores estivessem vinculados a instituições brasileiras, com o intuito de obter *e-mails* para contato – 572 *e-mails* identificados (Davies, 2008). Posteriormente esses *e-mails* foram analisados para eliminação de duplicidade.

   Além desses recursos, foram contactados o Conselho Federal de Medicina e sociedades profissionais, que não deram retorno.

Considerando os resultados obtidos no Brasil em relação ao total de profissionais existentes no país e por não ter conseguido atingir o número de respostas estatisticamente adequado de uma instituição, optou-se por prosseguir com a investigação apenas em Portugal (Tabela 61).

Tabela 61. **Demonstrativo de respostas no Brasil**

Brasil		
Questionários enviados	Questionários respondidos	*Links* respondidos
692	39	12

Do total de questionários enviados para o Brasil e Portugal, 63 foram devolvidos (problemas com identificação, endereço, *e-mail*) e 23 foram cancelados. Quanto à disseminação dos *links*, não foi possível determinar quantos tiveram acesso e deixaram de responder.

5.2.3 **O questionário**

A utilização de questionário é apontada por alguns autores como ferramenta que proporciona algumas vantagens em relação a outras técnicas de coleta de dados. No entanto, identifica-se na literatura que o número de respondentes é limitado (Fischbacher, Chappel, Edwards, & Summerton, 2000; VanGeest, Johnson, & Welch, 2007), principalmente no contexto médico. Dentre as outras opções de coleta de dados, há indícios de que o questionário, em vista da população estudada, parecia ser o mais adequado (Bryant, 2004, González-González et al., 2005). Considerou-se que dificuldades relativas à disponibilidade de tempo, à interferência do ambiente e às características inerentes aos investigados poderiam comprometer a utilização da entrevista ou da observação (Cheng, 2004, Woolf & Benson, 1989).

Como já foi apontado anteriormente, os profissionais da área da saúde lidam com vários fatores, que aqui foram denominados de categorias: diagnóstico, tratamento, epidemiologia e terapêutica, entre outros. Com o intuito de ser coerente com essa classificação, o questionário foi estruturado de acordo com essas categorias (P9-P17). As questões incluem resultados de estudos que identificaram TNIs de médicos ou investigadores em alguns domínios da Medicina (Tabelas 56 e 57), cujo resultado da categorização encontra-se na Tabela 58. Apenas um dos estudos analisados focou o domínio da Oncologia, precisamente em repositórios de amostras genéticas.

A utilização desses resultados buscou preencher a lacuna de informações sobre TNIs na Oncologia e incentivar os respondentes no processo de descoberta de outros TNIs para que pudessem preencher a questão aberta P17. Em adição, ainda é possível identificar quais TNIs apontadas por médicos e investigadores de outros domínios são transversais aos que atuam na Oncologia.

A partir desse entendimento, o questionário foi estruturado em cinco seções:

a) primeira, termo de consentimento;

b) segunda, dados demográficos;

c) terceira, identificação dos tipos de necessidades de informação em relação à descrição de artigos científicos (uma questão fechada de múltipla opção, com espaço para sugestões);

d) quarta, identificação dos TNIs, ou seja, informações específicas que podem ser obtidas ao buscar os artigos científicos (oito questões fechadas de múltipla escolha com opção para sugestões e uma aberta, para as quais não era exigida resposta). As questões constantes nessa seção foram desenvolvidas com base na Tabela 58, em que foram consideradas as categorias **Genética**,

Diagnóstico, **Terapêutica**, **Epidemiologia**, **Patologia**, **Imagiologia**, **Assistência à saúde** e **Paciente** e os termos a elas associados. Na questão aberta, foram utilizados campos para preenchimento de acordo com as seguintes categorias: Diagnóstico, Terapêutica, Epidemiologia, Patologia, Etiologia e Imagiologia, Procedimentos cirúrgicos, Compostos químicos e drogas, Técnicas de investigação, Equipamento e suplementos, sendo as quatro últimas extraídas da árvore hierárquica do MESH (https://www.ncbi.nlm.nih.gov/mesh/1000069).

e) a quinta e última seção, a conclusão (opção para identificação pessoal e comentários) (APÊNDICE C).

O questionário foi submetido a um pré-teste (piloto), analisado por três médicos (2 brasileiros e 1 português) e um investigador português, antes de ser submetido aos investigados (Grimm, 2010; Reynolds, Diamantopoulos, & Schlegelmilch, 1993). As sugestões de alteração provenientes dessa avaliação foram em relação à estrutura do questionário (tempo e número de questões) e à formulação do conteúdo (reduzir o número de TNIs e a questão aberta deveria estar alinhadas às categorias disponibilizadas para questões fechadas), tendo sido atendidas na sua quase totalidade. A redução do número de TNIs foi feita de forma a não prejudicar a identificação de TNIs – ênfase na eliminação da redundância e ambiguidade.

5.2.4 O resultado

Os dados deste estudo foram obtidos por meio do questionário eletrônico *SurveyMonkey* enviado para 301 investigados em Portugal, no período de julho de 2015 a maio de 2016. Cinquenta e nove questionários foram respondidos e os dados exportados para o Excel de forma a serem convertidos para o Statistical Package for the Social Sciences (SPSS), *software* de análise estatística. Após confirmar o vínculo profissional dos 59 respondentes, foram selecionados 39 correspondentes à amostra do Norte de Portugal, conforme justificativa já apresentada. Dos 59 respondentes, foram excluídos 4 questionários por não constar filiação ou a região onde os respondentes atuavam e 16 por pertencerem a outras regiões.

O tratamento dos dados possibilitou a identificação dos seguintes aspectos:

a) identificação demográfica dos respondentes – a idade predominante é inferior a 35 anos (Figura 24), o mestrado se destaca como formação, seguido do doutorado (Figura 25). Com esse resultado, a probabilidade de esse grupo ter tido a oportunidade de consultar algum tipo de serviço de informação *on-line* é grande, no entanto não é possível predizer o nível de familiaridade com esses serviços. Por

outro lado, é possível deduzir que os respondentes tiveram algum tipo de necessidade de informação vinculada, pelo menos, à consulta da literatura científica associada à obtenção do grau. A idade dos respondentes pode apontar para a facilidade que eles têm em interagir com a tecnologia.

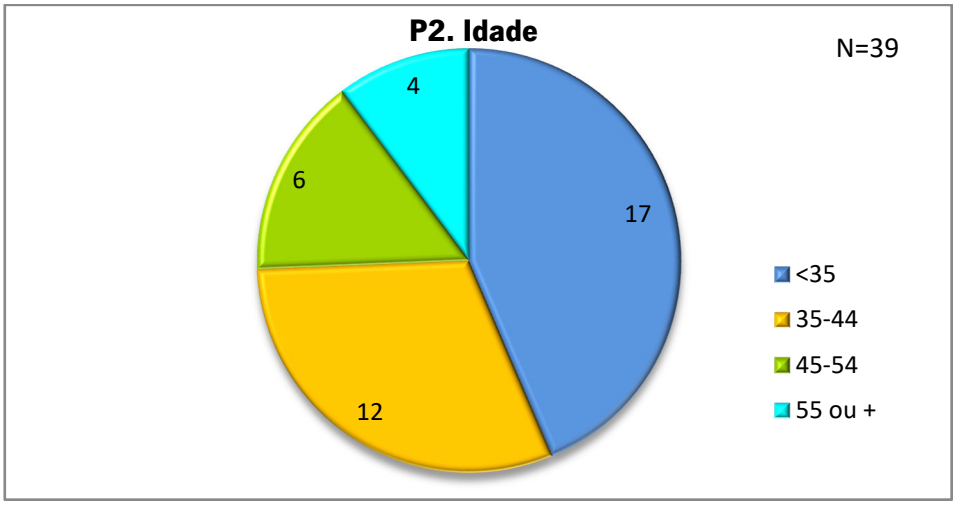


Figura 24. Idade dos respondentes.

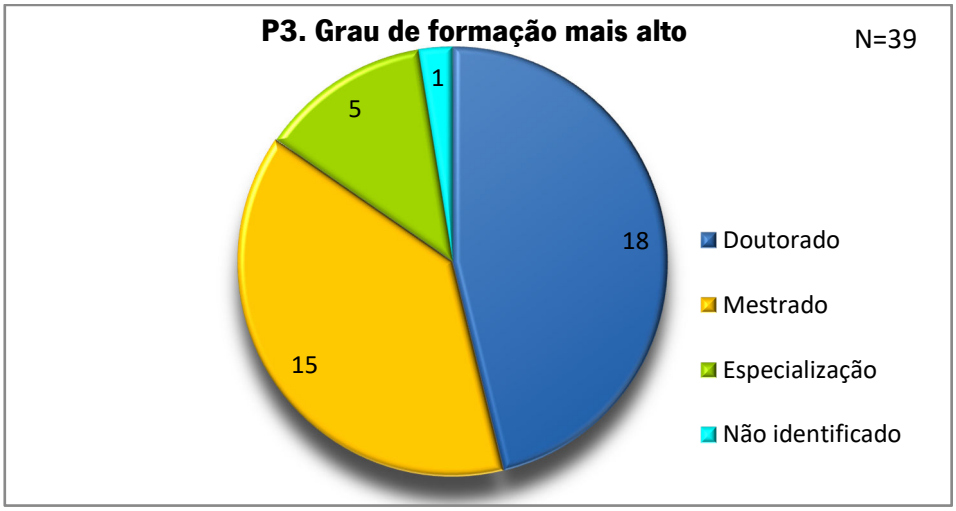


Figura 25. Formação dos respondentes.

O resultado apresentado na Figura 26, decorrente da participação de 39 respondentes, confirma as afirmações feitas pelos autores MacMullen & Denn (2005), Kim (2010b), Kim & Gilbertson (2007), Liverman et al. (1997), The Netherlands Cancer Institute Antoni van Leeuwenhock Hospital (2010), American Society of Clinical Oncology (2016) e Direcção-Geral da Saúde (2012) sobre a multidisciplinaridade no domínio da Oncologia.



Figura 26. Especialidade dos respondentes.

A Figura 27 apresenta as atividades desenvolvidas pelos 39 respondentes, em que se destaca o predomínio de investigadores, que pode ser explicado pelo vínculo empregatício desses respondentes com os institutos de investigação e com hospitais que também possuem centros de investigação ou colaboram com os institutos.

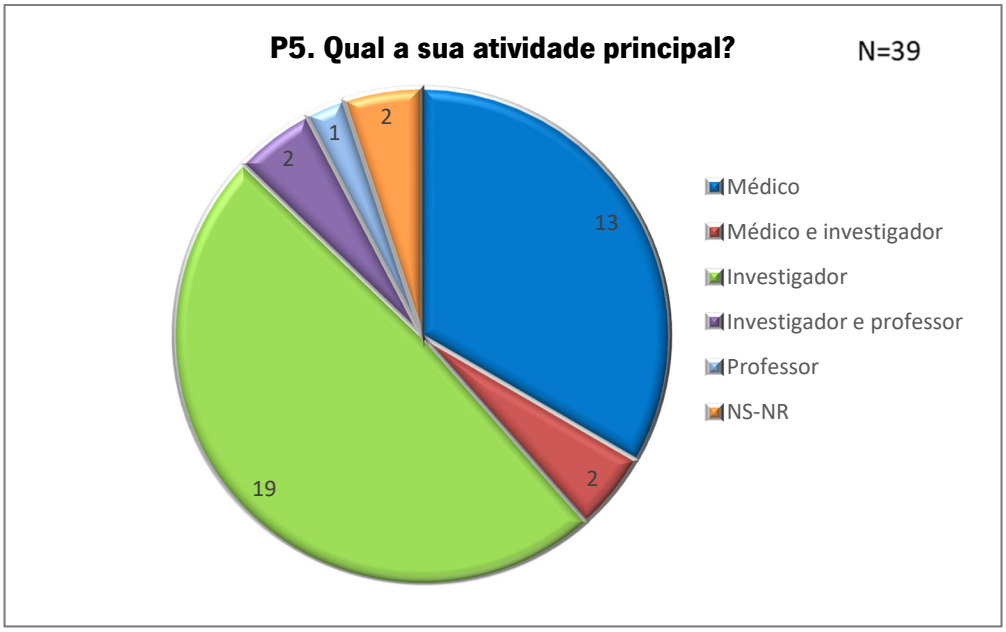


Figura 27. Atividades exercidas pelos respondentes.

O maior número de respondentes (19) afirma ter um período de experiência de 6 a 15 anos, o que pode indicar a familiaridade com questões que demandam informações especializadas para o atendimento clínico, assim como para o desenvolvimento de investigações científicas (Figura 28).

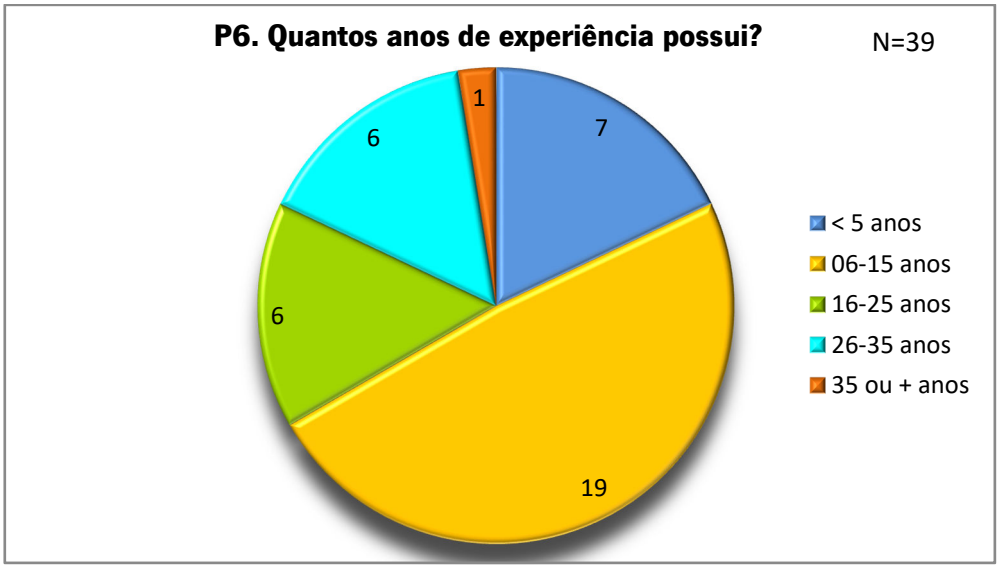


Figura 28. Anos de experiência dos respondentes.

O segundo aspecto tratado no *survey* corresponde à identificação dos TNIs (P8-17), que obteve os seguintes resultados:

c) Identificação dos TNIs – observa-se, na Figura 29, que os TNIs dos respondentes correspondentes a identificação de um recurso se referem predominantemente aos elementos tradicionais de busca, como: **título do artigo, assunto** (Bishop, 1998) e **autor**, mas também foram selecionados itens como **fator de impacto**, **figuras**, **tipos de estudos**, **notas** e **análise de citações**, resultado ratificado por Divoli, Wooldridge, & Hearst (2010). Esses elementos já se encontram disponíveis em algumas BDBs, conforme resultado da análise das bases apresentadas na Tabela 53. No entanto, ainda são pouco explorados pelos RDCs (Andrade & Baptista, 2015).

Dentre os elementos que podem ser utilizados para identificar um recurso ou partes dele, ainda é predominante o **título**, a **autoria** e o **assunto**. Entretanto, outros itens apontados na literatura consultada também são ratificados pelos respondentes como opção de busca, haja vista a Tabela 62, em que constam os dez itens mais citados.

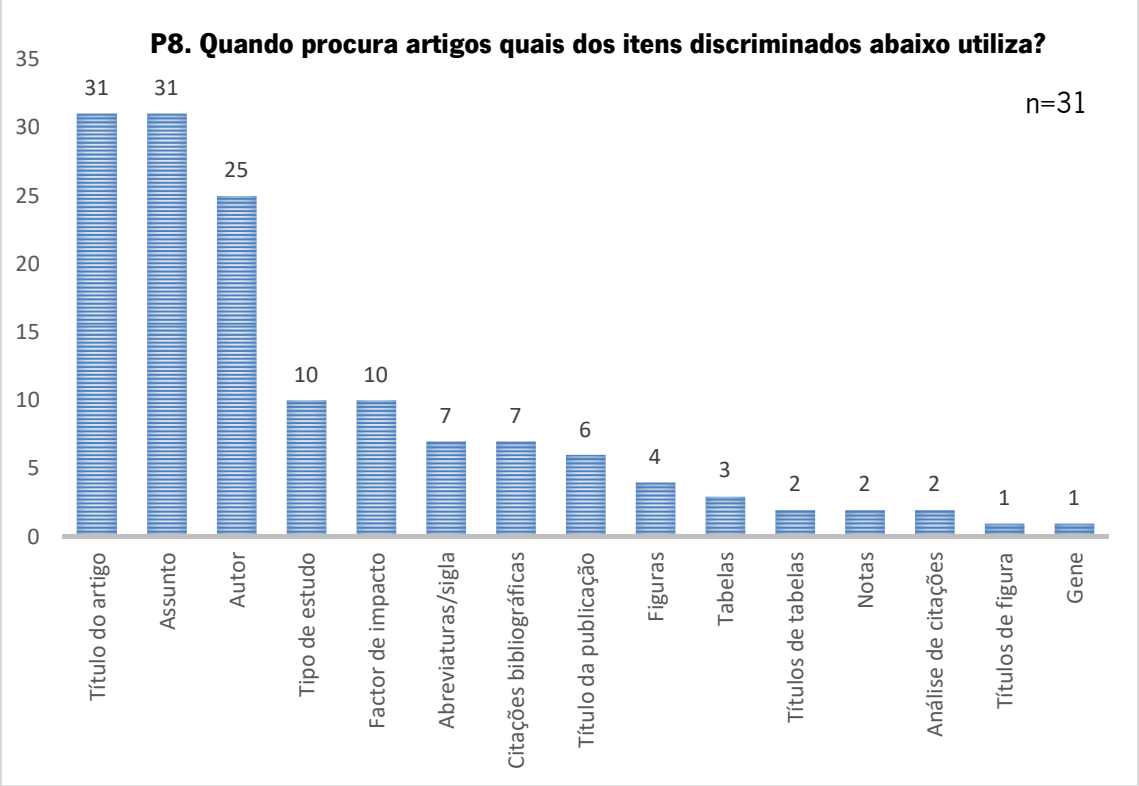


Figura 29. Itens utilizados pelos respondentes ao buscar artigos científicos.

Tabela. 62. **Itens utilizados com maior frequência pelos respondentes ao buscar artigos científicos**

Tipo de NI	Frequência	Percentagem
Titulo do artigo	31	79,5
Assunto	31	79,5
Autor	25	64,1
Tipo de estudo	10	25,6
Factor de impacto	10	25,6
Abreviaturas/siglas	7	17,9
Citações bibliográficas	7	17,9
Título da publicação	6	15,4
Figuras	4	10,3
Tabelas	3	7,7

Algumas tentativas de explicação para essas opções de busca podem ser encontradas no estudo de Hjørland, Nielsen, & Williams (2001). Os autores afirmam que, o que não é disponibilizado pelo sistema não pode ser utilizado. Esclarecem também que muitas vezes os utilizadores se restringem às informações disponibilizadas e, por não saberem da existência de outras opções, podem não percebê-las como TNIs.

Para as questões P9 a P17 não era obrigatória a resposta, visto que nem todos os profissionais necessitam do mesmo tipo de informação (Crowston & Kwasnik, 2003). Ou seja, nem todos os

participantes responderam a questão, como nem todos selecionaram todos os termos dela. Em vista da quantidade de TNIs incluídos nessas questões, às vezes apontados por um único respondente e com baixa frequência de uso, optou-se por apresentar apenas uma tabela com os dez TNIs mais indicados e com maior frequência de uso pelos respondentes, referentes a cada categoria.

A totalidade das propriedades e os índices de respostas são apresentados nos gráficos, em que foi omitida a quantidade referente aos que não responderam a questão.

Em relação à P9, dos 39 respondentes, 12 afirmaram ter TNIs em relação à Genética (Figuras 30 e 31).

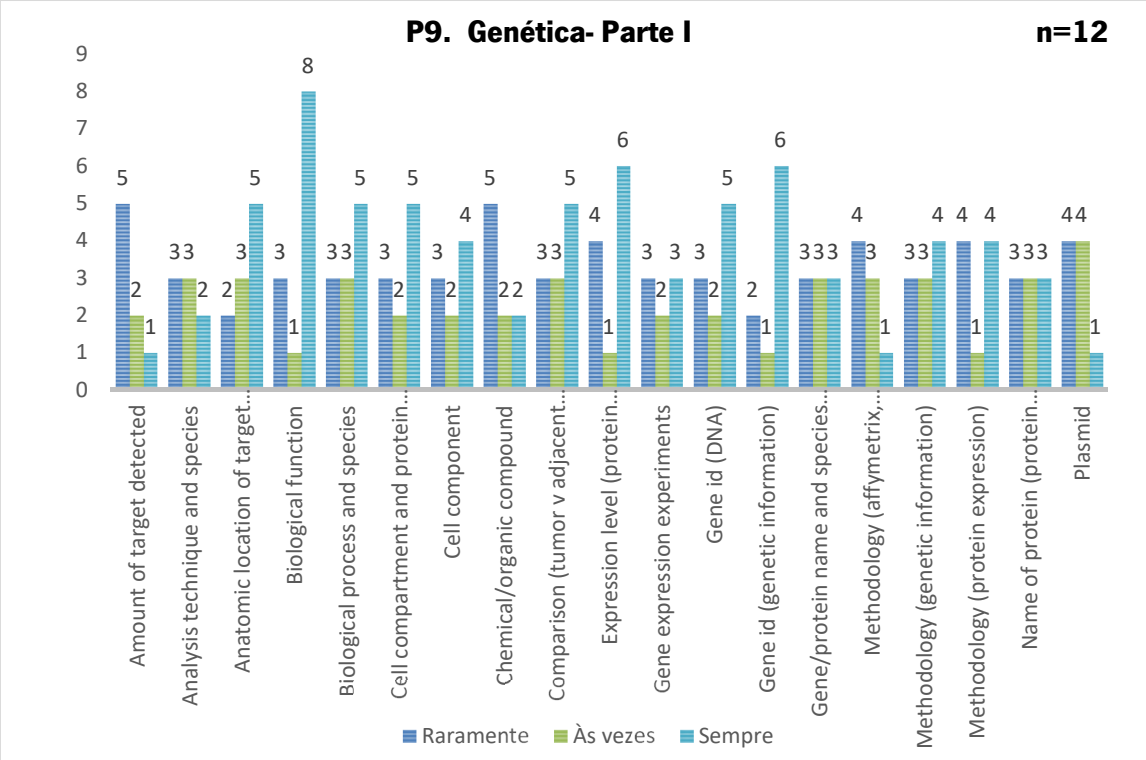


Figura 30. Tipos de NIs na categoria Genética – Parte 1.

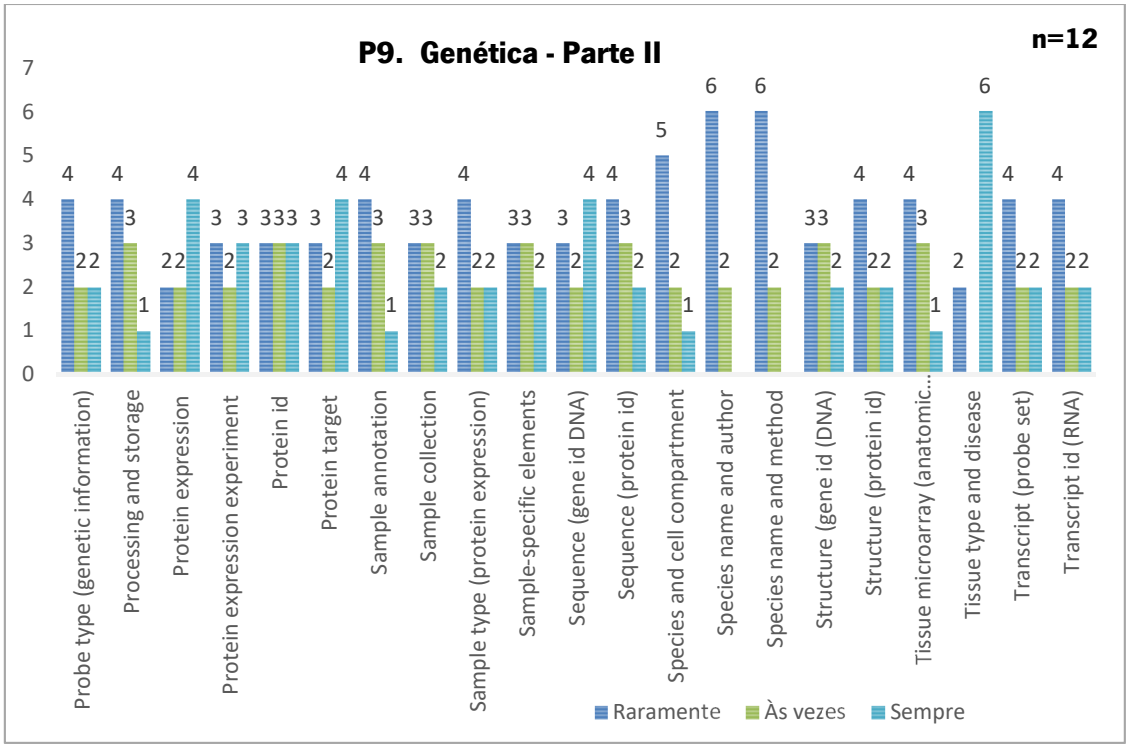


Figura 31. Tipos de NIs na categoria Genética – Parte 2.

Para os respondentes que buscam por Genética, foram identificados os dez TNIs mais utilizados pelos 12 respondentes (Tabela 63). Além de serem os dez tipos mais selecionados, também foram identificados como os utilizados com maior frequência.

Tabela 63. **Tipos de NIs mais frequentes na categoria Genética**

Tipo de NI - Genética	Raramente	Às vezes	Sempre	Não utiliza
Biological function (protein expression)	3 (7,7%)	1 (2,6%)	8 (20,5%)	27 (69,2%)
Expression level (protein expression)	4 (10,3%)	1 (2,6%)	6 (15,4%)	28 (71,7%)
Gene id	2 (5,1%)	1 (2,6%)	6 (15,4%)	30 (76,9%)
Tissue type and disease	2 (5,1%)	–	6 (15,4%)	31 (79,5%)
Comparison (tumor v adjacent normal)	3 (7,7%)	3 (7,7%)	5 (12,8%)	28 (71,8%)
Anatomic location of target detected	2 (5,1%)	3 (7,7%)	5 (12,8%)	29 (74,4%)
Cell compartment and protein structure	3 (7,7%)	2 (5,1%)	5 (12,8%)	29 (74,4%)
Cell component	3 (7,7%)	2 (5,1%)	4 (10,3%)	30 (76,9%)
Biological process and species	3 (7,7%)	3 (7,7%)	5 (12,8%)	28 (71,8%)
Gene id (DNA)	3 (7,7%)	2 (5,1%)	5 (12,8%)	29 (74,4%)

Em relação ao Diagnóstico (P10), observa-se que é uma das categorias com maior demanda por esses profissionais. Dos 39 entrevistados, 11 indicaram os TNIs apresentados nas Figuras 32 e 33. Os tipos com maior frequência de uso ao buscar artigos científicos podem ser visualizados na Tabela 64.

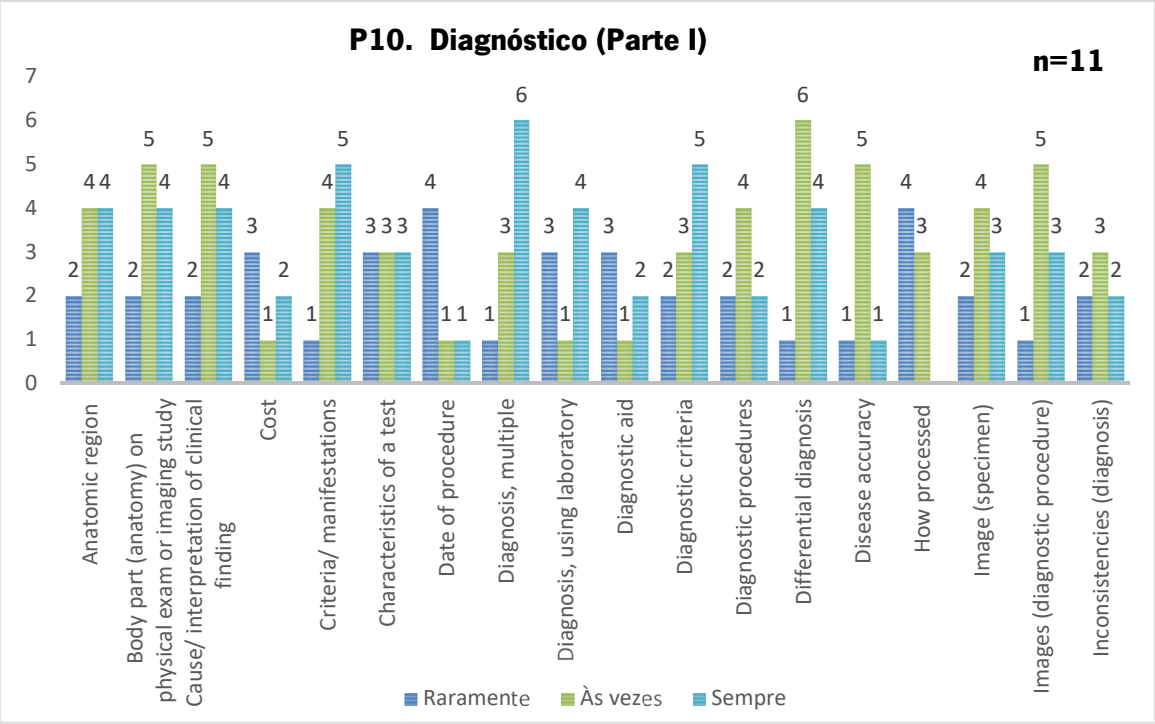


Figura 32. Tipos de NIs na categoria Diagnóstico – Parte 1.

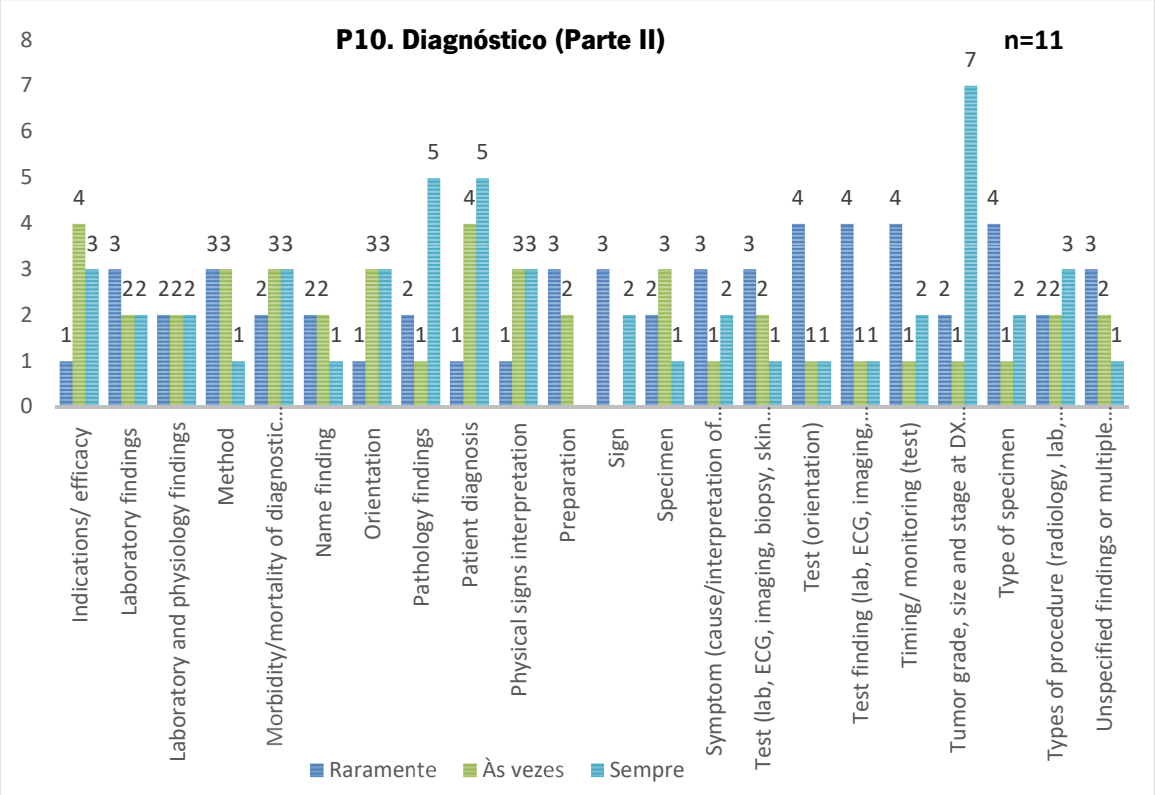


Figura 33. Tipos de NIs na categoria Diagnóstico – Parte 2.

Tabela 64. **Tipos de NIs mais frequentes na categoria Diagnóstico**

Tipo de NI: Diagnóstico	Raramente	Às vezes	Sempre	Não utilizam
Tumor grade, size and stage at DX and resection	2 (5,1%)	1 (2,6%)	7 (17,9%)	29 (74,4%)
Diagnosis, multiple	1 (2,6%)	3 (7,7%)	6 (15,4%)	29 (74,4%)
Criteria/ manifestations	1 (2,6%)	4 (10,3%)	5 (12,8%)	29 (74,4%)
Diagnostic criteria	2 (5,1%)	3 (7,7%)	5 (12,8%)	29 (74,4%)
Pathology findings (specimen)	2 (5,1%)	1 (2,6%)	5 (12,8%)	31 (79,5%)
Patient diagnosis	1 (2,6%)	4 (10,3%)	5 (12,8%)	29 (74,4%)
Body part (anatomy) on physical exam or imaging study	2 (5,1%)	5 (12,8%)	4 (10,4%)	28 (71,8%)
Anatomic region (specimen)	2 (5,1%)	4 (10,3%)	4 (10,3%)	29 (74,4%)
Cause/ interpretation of clinical finding	2 (5,1%)	5 (12,8%)	4 (10,3%)	11 (71,8%)
Diagnosis, using laboratory	3 (7,7%)	1 (2,6%)	4 (10,3%)	31 (79,5%)

A Questão P11, que aborda a categoria Terapêutica, teve 9 respondentes dos 39 participantes. Os termos selecionados pelos respondentes são apresentados nas Figuras 34 e 35.

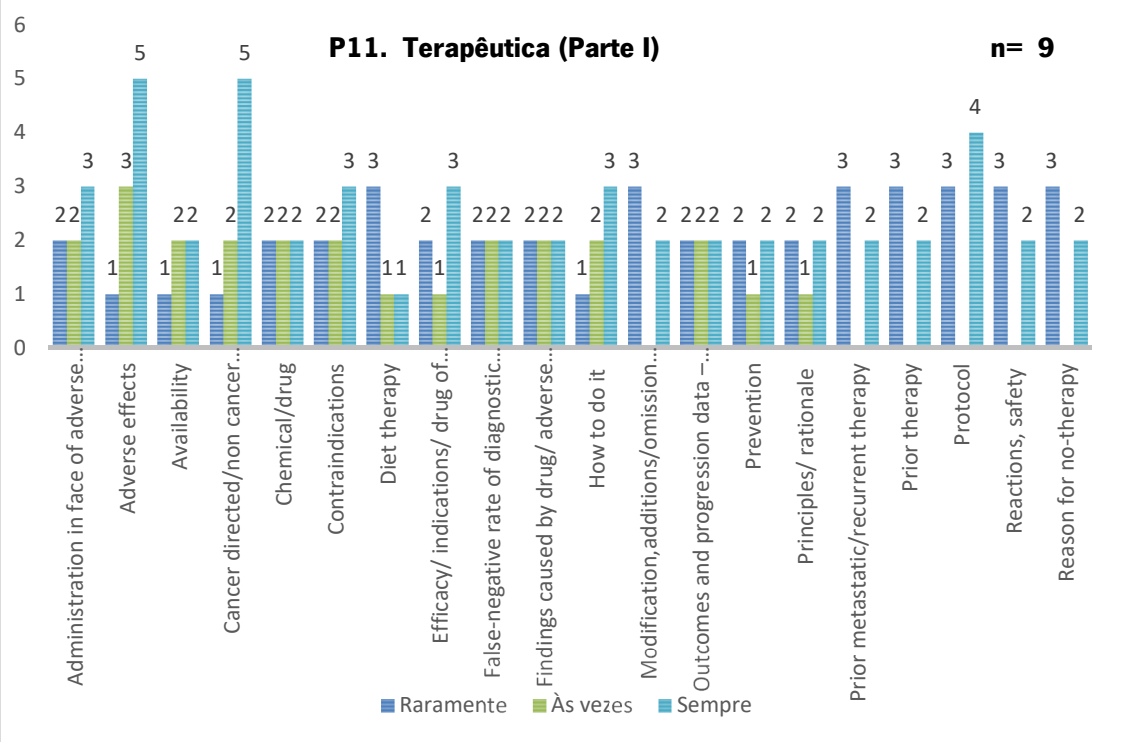


Figura 34. Tipos de NIs na categoria Terapêutica – Parte 1.

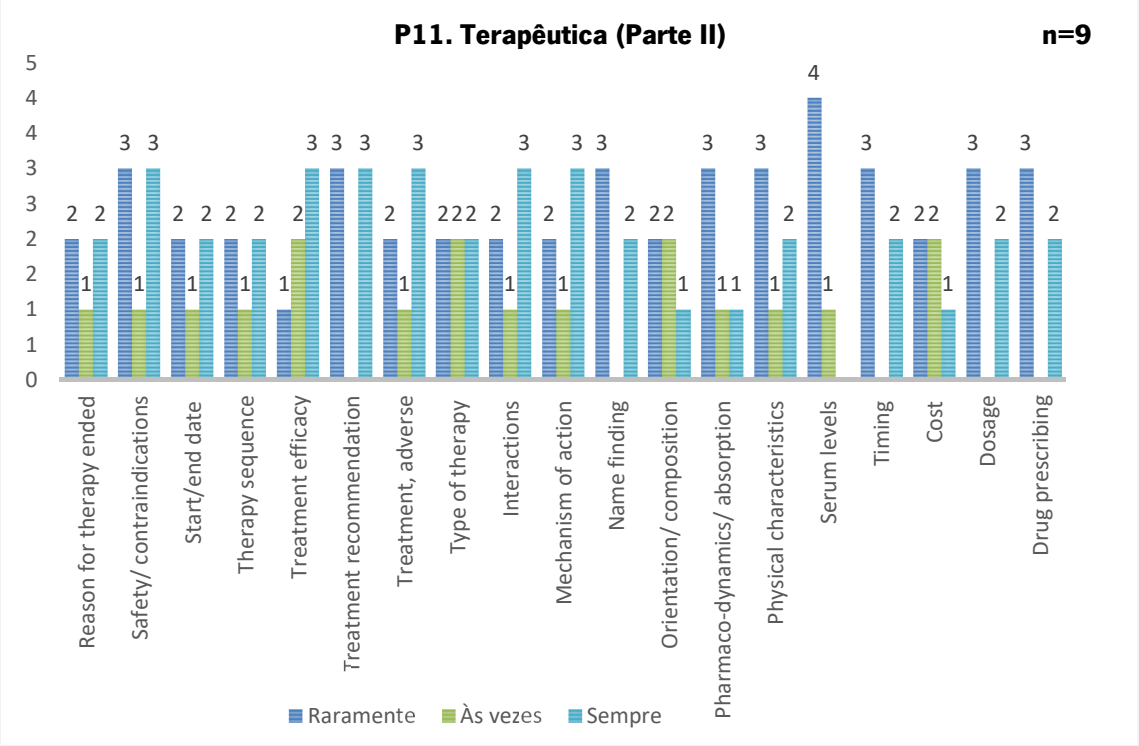


Figura 35. Tipos de NIs na categoria Terapêutica – Parte 2.

Dos termos selecionados pelos entrevistados que responderam à questão, foram selecionados os dez termos utilizados com maior frequência no processo de busca (Tabela 65).

Tabela 65. **Tipos de NIs mais frequentes na categoria Terapêutica**

Tipo de NI: Terapêutica	Raramente	Às vezes	Sempre	Não utiliza
Adverse effects	1 (2,6%)	3 (7,7%)	5 (12,8%)	9 (76,9%)
Cancer directed/non cancer directed	1 (2,6%)	2 (5,1%)	5 (12,8%)	31 (79,5)
Modification,additions/omission to therapy	3 (7,7%)		5 (12,8%)	34 (87,2%)
Dosage	3 (7,7%)		2(5,1%)	34 (87,2%)
Protocol (therapeutic procedures)	3 (7,7%)		4 (10,3%)	32 (82,1%)
Administration in face of adverse effects	2 (5,1%)	2 (5,1%)	3 (7,7%)	32 (82,1%)
Contraindications	2 (5,1%)	2 (5,1%)	3 (7,7%)	32 (82,1%)
Efficacy/ indications/ drug of choice	2 (5,1%)	1 (2,6%)	3 (7,7%)	33 (84,6%)
How to do it (treatment)	1 (2,6%)	2 (5,1%)	3 (7,7%)	33 (84,6%)
Safety/ contraindications (includes pregnancy and breast feeding)	3 (7,7%)	1 (2,6%)	3 (7,7%)	32 (82,1%)

Com referência à Questão P12, sobre Epidemiologia, dos 39 entrevistados, 9 responderam. O número de variáveis incluídas nesta questão foi menor do que o das demais questões (Figura 36). Esse número de variáveis foi mantido para que fosse seguido o mesmo procedimento das outras questões

– incluir elementos identificados na revisão de literatura. Nota-se que esses termos, quando utilizados, são selecionados com maior frequência (Tabela 66).

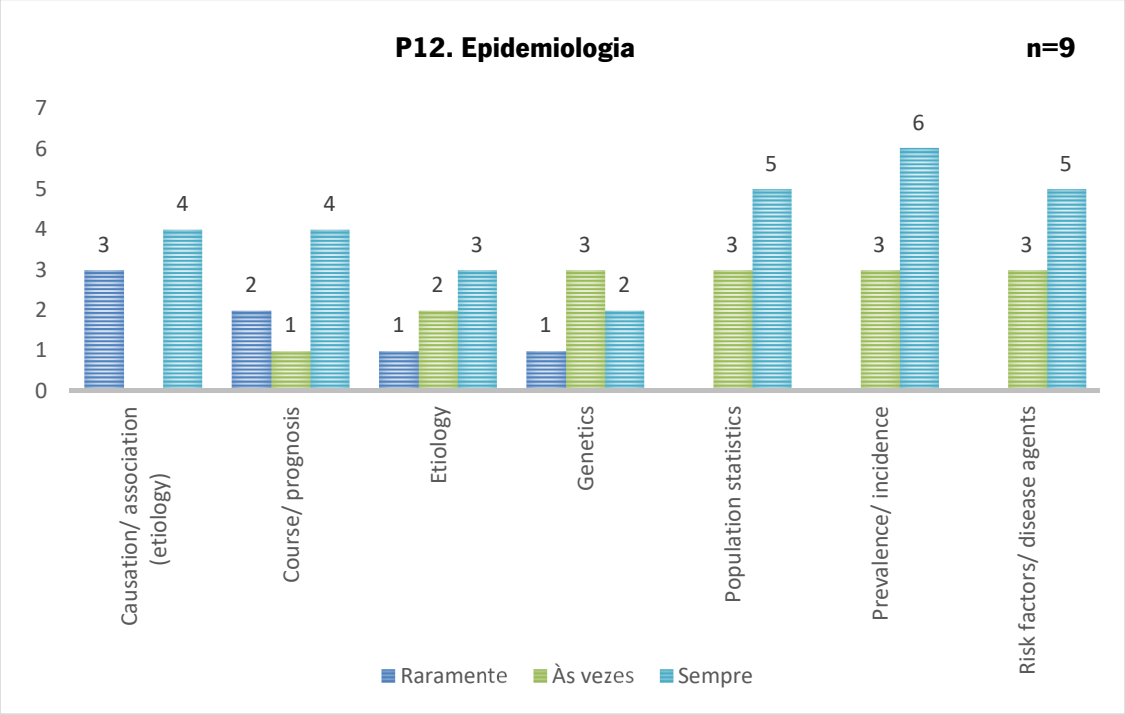


Figura 36. Tipos de NIs na categoria Epidemiologia.

Tabela 66. **Tipos de NIs mais frequentes na categoria Epidemiologia**

Tipo de NI: Epidemiologia	Raramente	Às vezes	Sempre	Não utiliza
Prevalence/ incidence		3 (7,7%)	6 (15,4%)	30 (76,9%)
Population statistics		3 (7,7%)	5 (12,8%)	31 (79,5%)
Risk factors/ disease agents		3 (7,7%)	5 (12,8%)	31 (79,5%)
Causation/ association (etiology)	3 (7,7%)		4 (10,3%)	32 (82,1%)
Course/ prognosis	2 (5,1%)	1 (2,6%)	4 (10,3%)	32 (82,1%)
Etiology	1 (2,6%)	2 (5,1%)	3 (7,7%)	33 (84,6%)
Genetics	1 (2,6%)	3 (7,7%)	2 (5,1%)	33 (84,6%)

Na P13, TNIs inerentes à Patologia, 14 dos 39 respondentes sinalizaram que todos os tipos apresentados são utilizados (Figura 37). Os dez tipos mais utilizados e com maior frequência são expostos na Tabela 67. Esta questão também revelou maior número de respondentes.

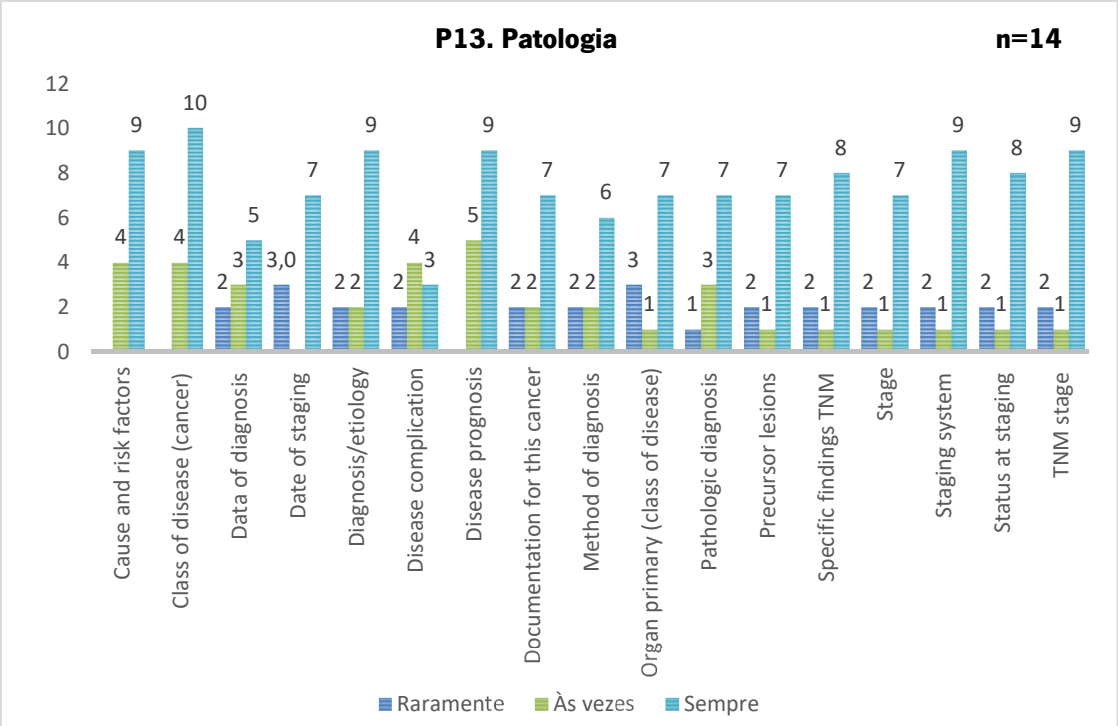


Figura 37. Tipos de NIs na categoria Patologia.

Tabela 67. **Tipos de NIs mais frequentes na categoria Patologia**

Tipo de NI: Patologia	Raramente	Às vezes	Sempre	Não utiliza
Class of disease (cancer) (nature of disease)		4 (10,3%)	10 (25,6%)	25 (64,1%)
Cause and risk factors		4 (10,3%)	9 (23,1%)	26 (66,7%)
Diagnosis/etiology	2 (5,1%)	2 (5,1%)	9 (23,1%)	26 (66,7%)
Disease prognosis		5 (12,8%)	9 (23,1%)	25 (64,1%)
Staging system (stage) (extent of disease)	2 (5,1%)	1 (2,6%)	9 (23,1%)	27 (69,2%)
TNM stage (extent of disease)	2 (5,1%)	1 (2,6%)	9 (23,1%)	27 (69,2%)
Specific findings TNM (stage) (extent of disease)	2 (5,1%)	1 (2,6%)	8 (20,5%)	28 (71,8%)
Status at staging (extent of disease)	2 (5,1%)	1 (2,6%)	8 (20,5%)	28 (71,8%)
Date of staging (extent of disease)	3 (7,7%)		7 (17,9%)	29 (74,4%)
Documentation for this cancer) (nature of disease)	2 (5,1%)	2 (5,1%)	7 (17,9%)	28 (71,8%)

Houve um baixo índice de respondentes para o uso dos TNIs (n=7) apresentados na categoria Imagiologia (P14), principalmente se analisados na perspectiva daqueles que são sempre utilizados (Figuras 38 e 39). Talvez uma possível explicação seja o número reduzido de radio-oncologistas (2) e pelos TNIs incluídos no questionário estarem relacionados com aspectos técnicos da imagem. Os

itens utilizados com maior frequência pelos que responderam à questão estão explicitados na Tabela 68.

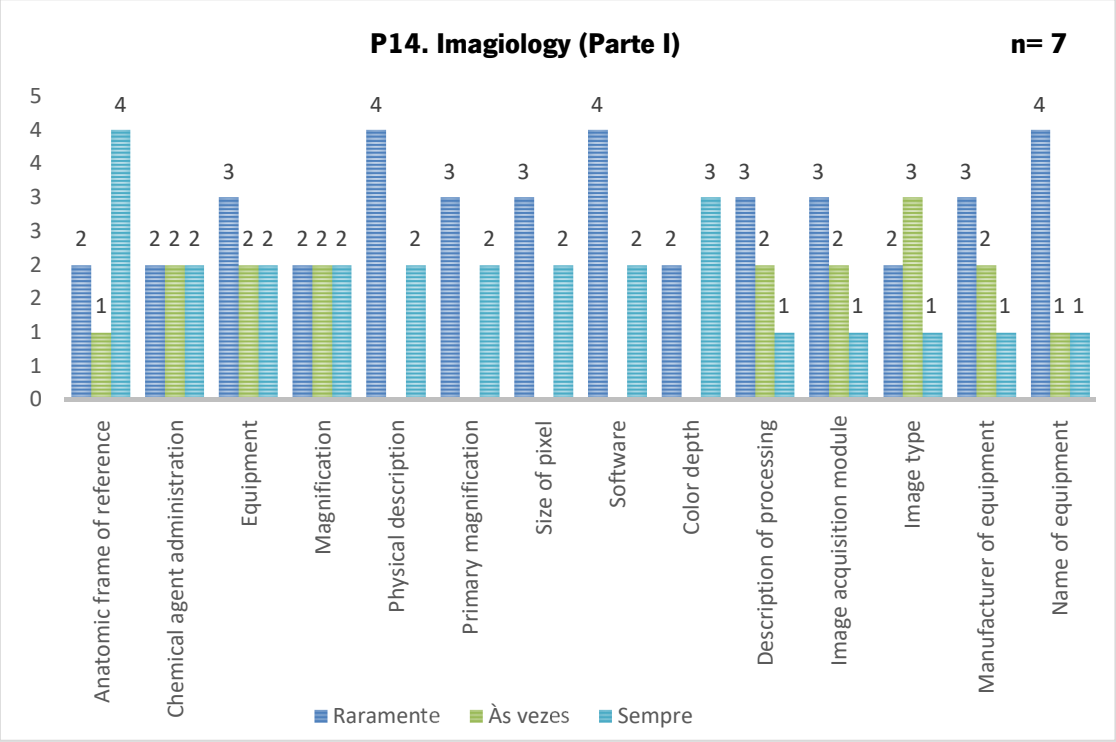


Figura 38. Tipos de NIs na categoria Imagiologia – Parte 1.

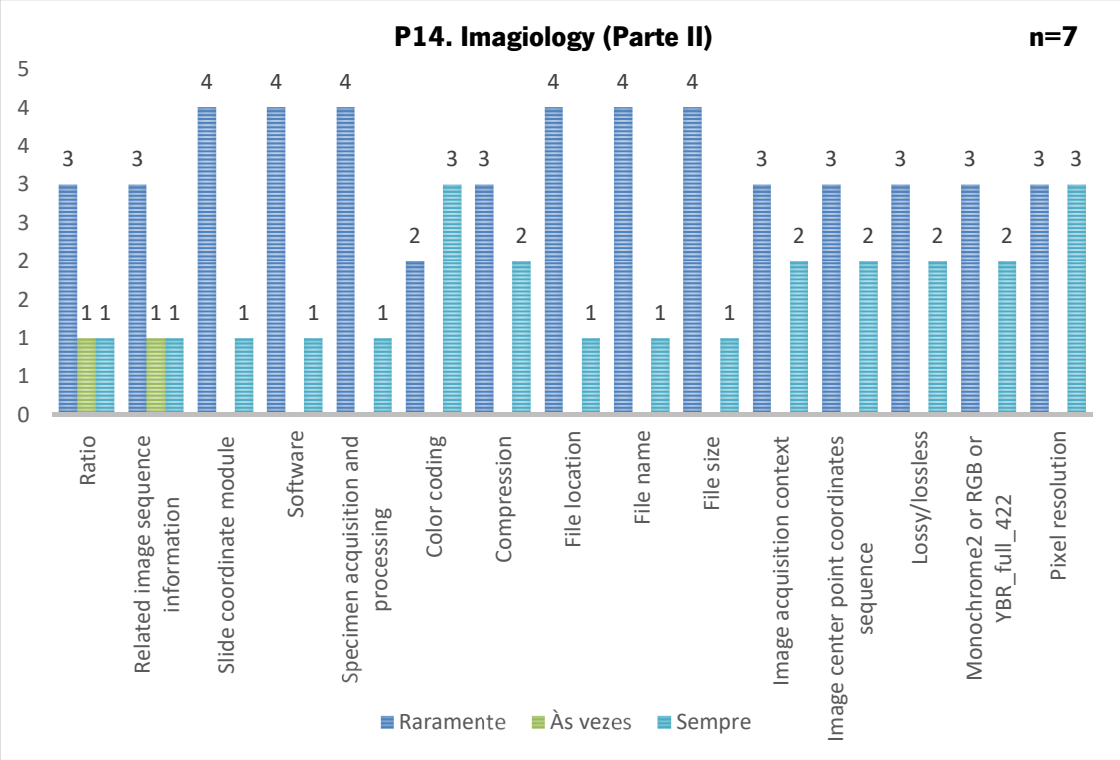


Figura 39. Tipos de NIs na categoria Imagiologia – Parte 2.

Tabela 68. **Tipos de NIs mais frequentes na categoria Imagiologia**

Tipo de NI: Imagiologia	Raramente	Às vezes	Sempre	Não utiliza
Anatomic frame of reference	2 (5,1%)	1 (2,6%)	4 (10,3%)	32 (82,1%)
Chemical agent administration	2 (5,1%)	2 (5,1%)	2 (5,1%)	33 (84,6%)
Equipment	3 (7,7%)	2 (5,1%)	2 (5,1%)	32 (82,1%)
Magnification	2 (5,1%)	2 (5,1%)	2 (5,1%)	33 (84,6%)
Physical description	4 (10,3%)	—	2 (5,1%)	33 (84,6%)
Primary magnification	3 (7,7%)	—	2 (5,1%)	34 (87,2%)
Size of pixel	3 (7,7%)	—	2 (5,1%)	34 (87,2%)
Software	4 (10,3%)	—	2 (5,1%)	33 (84,6%)
Color depth	2 (5,1%)	3 (7,7%)	1 (2,6%)	33 (84,6%)
Description of processing	3 (7,7%)	2 (5,1%)	1 (2,6%)	33 (84,6%)

A assistência à saúde corresponde à Questão P15, em que sete respondentes indicaram 11 TNIs (Figura 40).

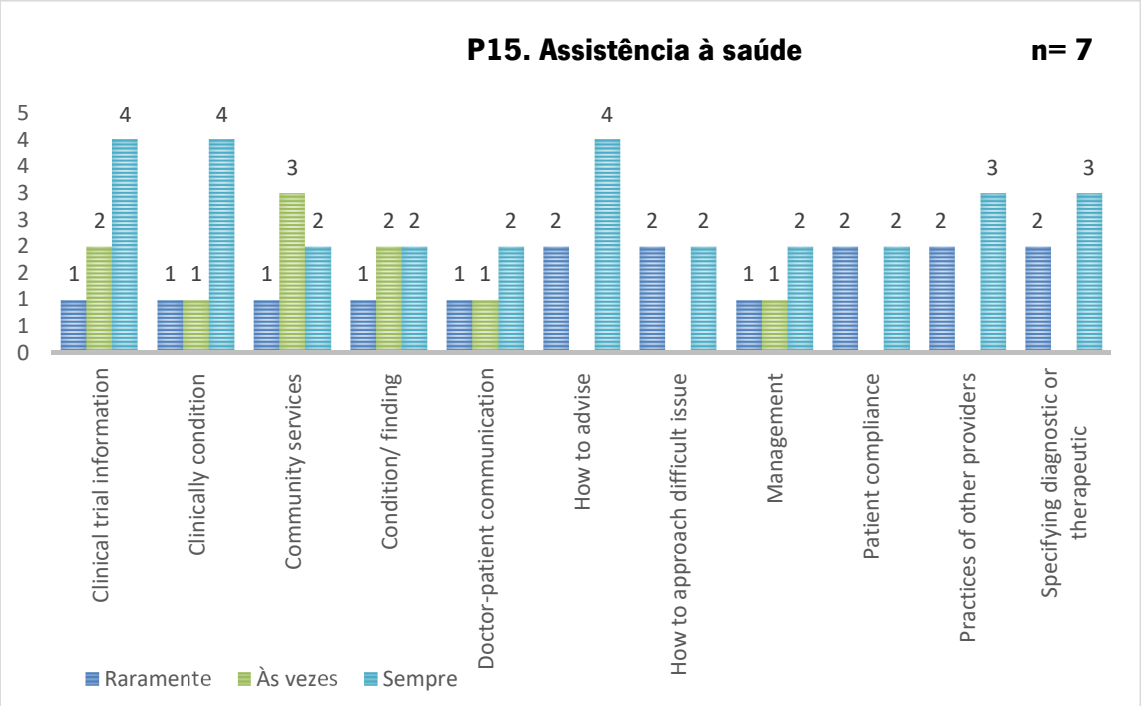


Figura 40. Tipos de NIs na categoria Assistência à saúde.

A Tabela 69 apresenta os dez TNIs mais utilizados por ordem de frequência de uso.

Tabela 69. **Tipos de NIs mais frequentes na categoria Assistência à Saúde**

Tipo de NI: Assistência à saúde	Raramente	Às vezes	Sempre	Não utiliza
Clinical trial information	1 (2,6%)	2 (5,1%)	4 (10,3%)	32 (82,1%)
Clinically condition (orientation)	1 (2,6%)	1 (2,6%)	4 (10,3%)	33 (84,6%)
Community services (management)	1 (2,6%)	3 (7,7%)	2 (5,1%)	33 (84,6%)
Condition/ finding (management)	1 (2,6%)	2 (5,1%)	2 (5,1%)	34 (87,2%)
Doctor-patient communication	1 (2,6%)	1 (2,6%)	2 (5,1%)	35 (89,7%)
How to advise (management)	2 (5,1%)	—	4 (10,3%)	33 (84,6%)
How to approach difficult issue (doctor-patient communication)	2 (5,1%)	—	2 (5,1%)	35 (89,7%)
Management (not specifying diagnostic or therapeutic) (management)	1 (2,6%)	1 (2,6%)	2 (5,1%)	35 (89,7%)
Patient compliance (doctor-patient communication)	2 (5,1%)		2 (5,1%)	35 (89,7%)
Practices of other providers (management)	2 (5,1%)		3 (7,7%)	34 (87,2%)

O acompanhamento do paciente oncológico requer dos médicos e investigadores informações em relação aos TNIs elencados na Figura 41. Dentre esses tipos, foram selecionados os dez tipos com maior frequência de uso pelos sete respondentes (Tabela 70).

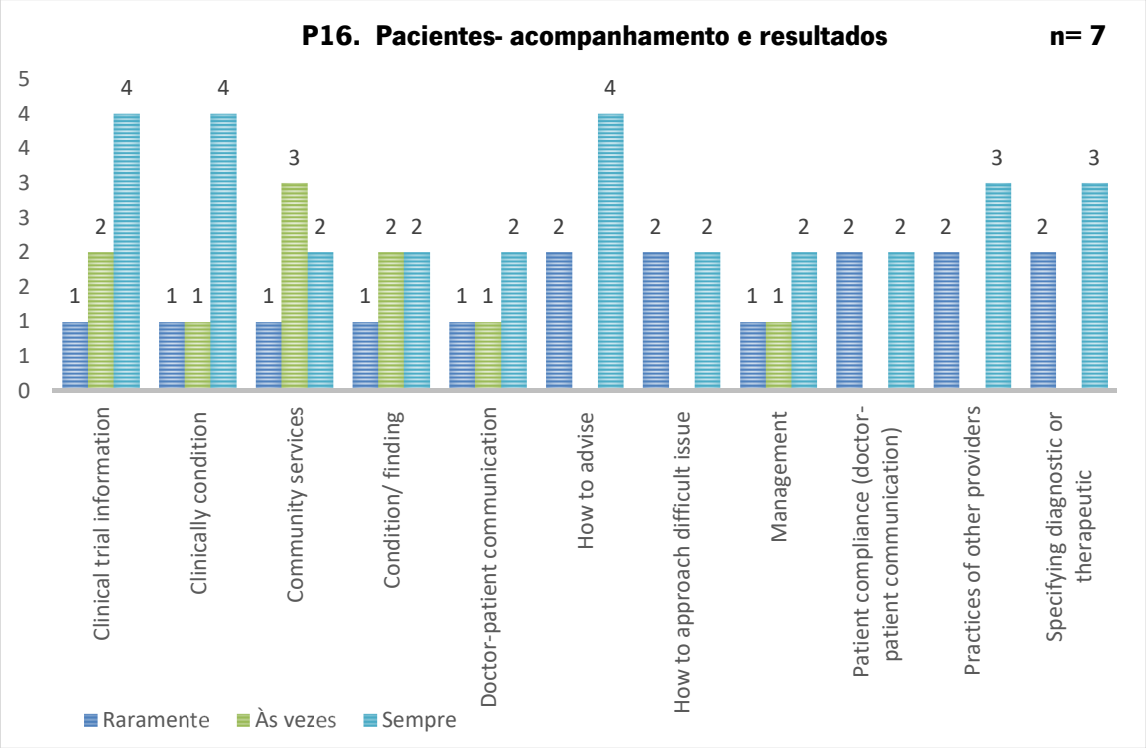


Figura 41. Tipos de NIs na categoria Pacientes.

Tabela 70. **Tipos de NIs mais frequentes na categoria Pacientes**

Tipos de NIs:Pacientes-acompanhamento e resultados	Raramente	Às vezes	Sempre	Não utiliza
Site of metastasis	2 (5,1%)	1 (2,6%)	7 (17,9%)	29 (74,4%)
Family history of disease	1 (2,6%)	3 (7,7%)	6 (15,4%)	29 (74,4%)
Metastases (multiple)	2 (5,1%)	1 (2,6%)	6 (15,4%)	30 (76,9%)
Primary cause of death	2 (5,1%)	1 (2,6%)	6 (15,4%)	30 (76,9%)
How documented (biopsy, cytology, radiology exam)	2 (5,1%)		5 (12,8%)	32 (82,1%)
How a patient should be managed disease information	2 (5,1%)		5 (12,8%)	32 (82,1%)
Method used to determine first recurrence	2 (5,1%)		5 (12,8%)	32 (82,1%)
Type of first recurrence (local, distant, biochemical)	2 (5,1%)		5 (12,8%)	32 (82,1%)
Vital status	2 (5,1%)		5 (12,8%)	32 (82,1%)
Cancer status at death	2 (5,1%)	2 (5,1%)	4 (10,3%)	31 (79,5%)

A penúltima questão (P17) foi projetada com o objetivo de identificar outros TNIs não apresentados nas questões anteriores e, assim, proporcionar uma visão mais particular dos profissionais que atuam na Oncologia. Houve pouca manifestação ou redundância quanto aos tipos já disponibilizados. O número reduzido de respostas para essa questão é justificado por Fourie & Baker (2009), ao afirmarem que os profissionais da saúde não expandem muito as suas necessidades de informação (Fourie & Baker, 2009). As inclusões foram realizadas em relação às categorias:

a) Diagnóstico: ***biomarker***;

b) Terapêutica: terapêuticas utilizadas atualmente;

c) Procedimentos cirúrgicos: *what was the procedure carried out*;

d) Epidemiologia: **incidência**, incidência por tipo de câncer;

e) Compostos químicos e drogas: **efeitos biológicos e mecanismo e acção**, **efeitos adversos**;

f) Patologia: *what previous diagnosis have been established*, ***disease***, **mecanismo biológico**;

g) Etiologia: **radiation**;

h) Imagiologia: *what is the location and characteristics of the present lesion*, **radioiodine**;

i) Técnicas de investigação (utilizada na pré-clínica, investigação clínica, epidemiologia, imunologia e genética): crisp;

j) Equipamento e suplementos.

A questão P18 possibilitava a inclusão de novas categorias ou TNIs. Como resultado, foram apontados os termos *Biomarker* e *What is the clinical diagnosis*.

Evidencia-se, a partir desses resultados, que as categorias diagnóstico e terapêutica concentraram maior número de respondentes em relação a outras categorias, o que é corroborado pela literatura consultada.

Os dez TNIs com maior frequência de uso referente a cada categoria foram explicitados em um gráfico denominado de Taxonomia dos tipos de necessidade de informação em Oncologia, representado na Figura 42.

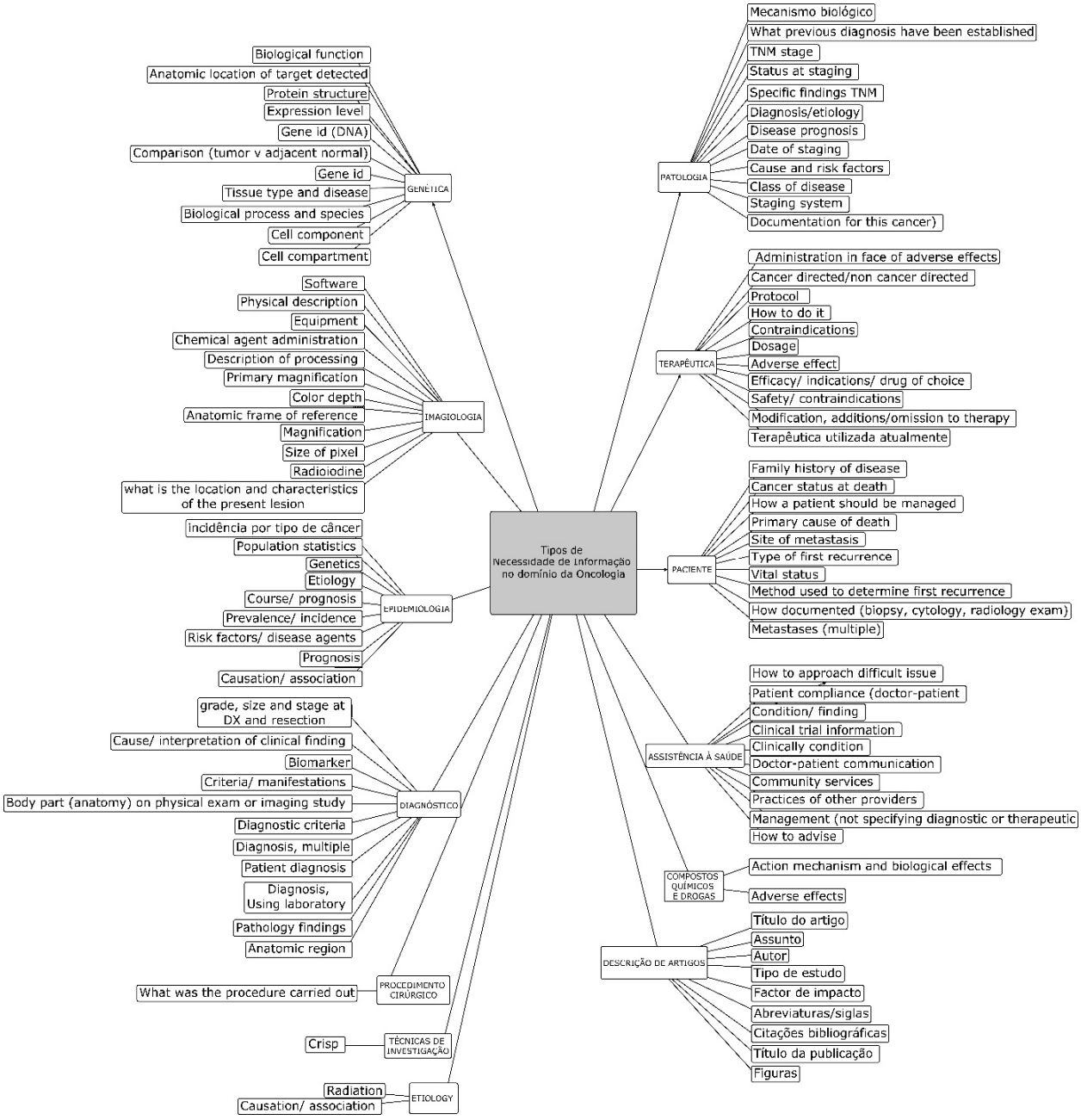


Figura 42. Taxonomia dos tipos de necessidade de informação em Oncologia.

SÍNTESE DA SECÇÃO 5

Esta Seção apresenta a investigação desenvolvida para identificação dos TNIs dos médicos e investigadores que atuam na Oncologia. Foram detalhados os procedimentos adotados na pesquisa bibliográfica e o estudo de caso com médicos e investigadores que atuam em Portugal no domínio da Oncologia. No estudo de caso, foi adotado o Survey, que é detalhado em relação a forma de acesso aos investigados, construção do instrumento de coleta de dados (questionário), aplicação do inquérito e análise dos resultados. A investigação resultou na seleção de 101 TNIs, com base no número de indicações e de frequência de uso pelos respondentes, referentes a 13 categorias.

A organização das informações resultantes da investigação originou a Taxonomia dos tipos de necessidade de informação em Oncologia.

6 COMPILAÇÃO DOS RESULTADOS E O CICLO DA RELEVÂNCIA DE HEVNER

Este capítulo reúne os resultados dos estudos desenvolvidos no âmbito desta tese, identificados no Design Science Research (DSR) de Hevner (Seção 3), como a identificação de requisitos no Ciclo da Relevância. Na dinâmica do DSR, esse ciclo estabelece a ponte entre a Fase Ambiente e a Fase do Desenvolvimento do Artefacto, ou seja, a fase em que o artefacto é desenvolvido e validado (Figura 15).

Assim, a proposta deste capítulo, é apresentar, além dos elementos que integrarão o PAM (Fase Ambiente, Ciclo da Relevância), o resumo dos procedimentos adotados para identificá-los, associados à Base de Conhecimento e ao Ciclo do Rigor.

6.1 SELEÇÃO DOS TIPOS DE NECESSIDADES DE INFORMAÇÃO

Os TNIs foram selecionados a partir dos estudos realizados no contexto desta tese, os quais estão descritos sucintamente a seguir.

Estudo I – Identificação de PAMs, esquemas e elementos de metadados

Objetivo: localizar e analisar a literatura com o objetivo de identificar PAMs, esquemas e elementos de metadados.

Atividade: analisar de forma sistemática artigos, teses, dissertações, *papers* e documentação obtidos por meio de pesquisas em bases de dados bibliográficas, *sites* de instituições, repositórios digitais científicos e motores de busca (Google Scholar).

Resultado: identificar e categorizar os elementos de metadados descritos nos PAMs.

Atividade complementar: *Survey* Application Profile and Metadata Schemas – complementação à pesquisa documental sobre PAMs. Foram identificados PAMs adotados por RDCs; características dos RDCs, conteúdo armazenado e esquemas de metadados adotados (Figura 43).

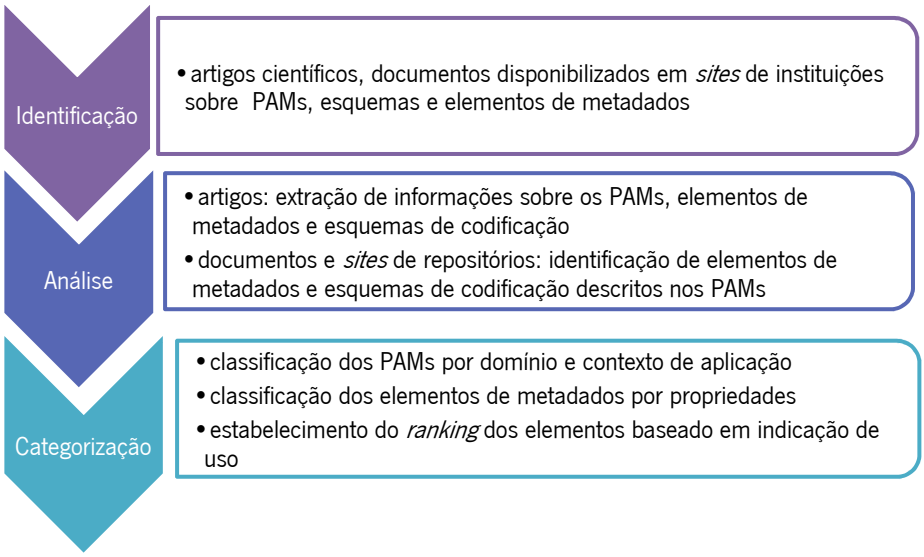


Figura 43. Pesquisa documental sobre perfis de aplicação, esquemas e elementos de metadados.

Estudo II – Análise da estrutura de 13 bases de dados bibliográficas que cobrem o domínio da Oncologia

Objetivo: identificar os pontos de acesso de bases de dados bibliográficas

Atividades apresentadas na Figura 44:

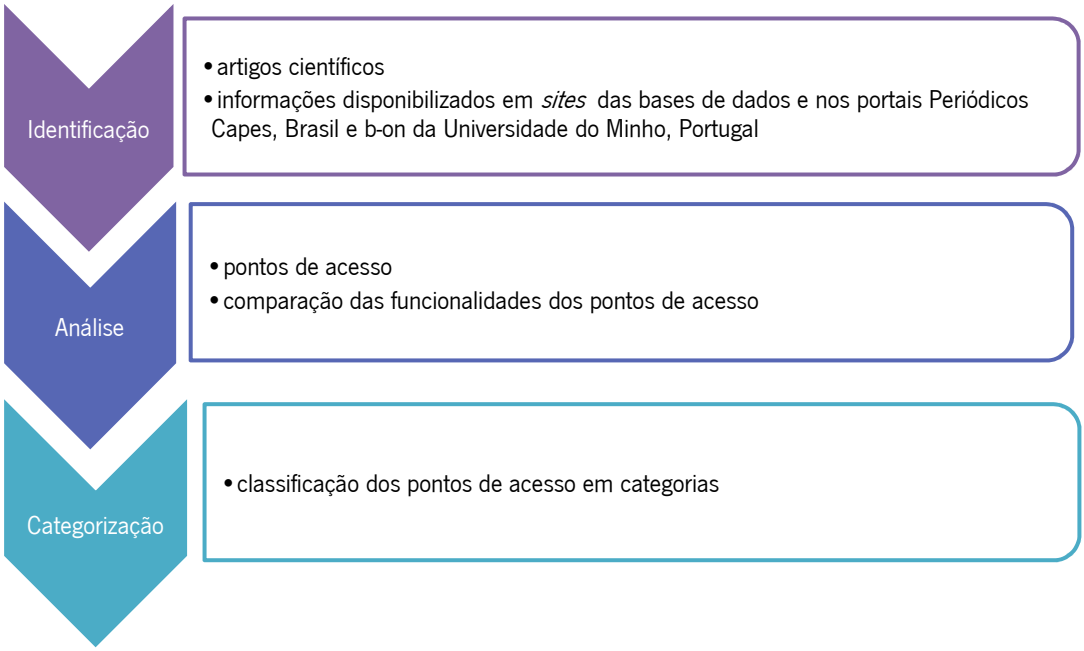


Figura 44. Pesquisa documental sobre pontos de acesso de bases de dados bibliográficas.

A autora, por inferência, entende que esses resultados podem representar TNIs. A autora considerou informações sobre o desenvolvimento de PAMs com base nas necessidades dos seus utilizadores e a sua experiência com investigadores em relação à identificação de artigos a partir dos pontos de acesso.

Estudo III – Estudo de caso: Tipos de NIs de médicos e investigadores no domínio da Oncologia em Portugal

Objetivo: Identificar os TNIs de médicos e investigadores que atuam em Oncologia.

Atividades:

a) analisar a literatura sobre necessidade informacional de médicos e investigadores no domínio da Medicina;

b) aplicar *Survey* com envio do questionário (*Survey*Monkey) e divulgação do *link* para médicos e investigadores que atuam em Portugal;

c) selecionar TNIs de cada categoria baseada no maior número de utilização e maior frequência de uso (Figura 45).

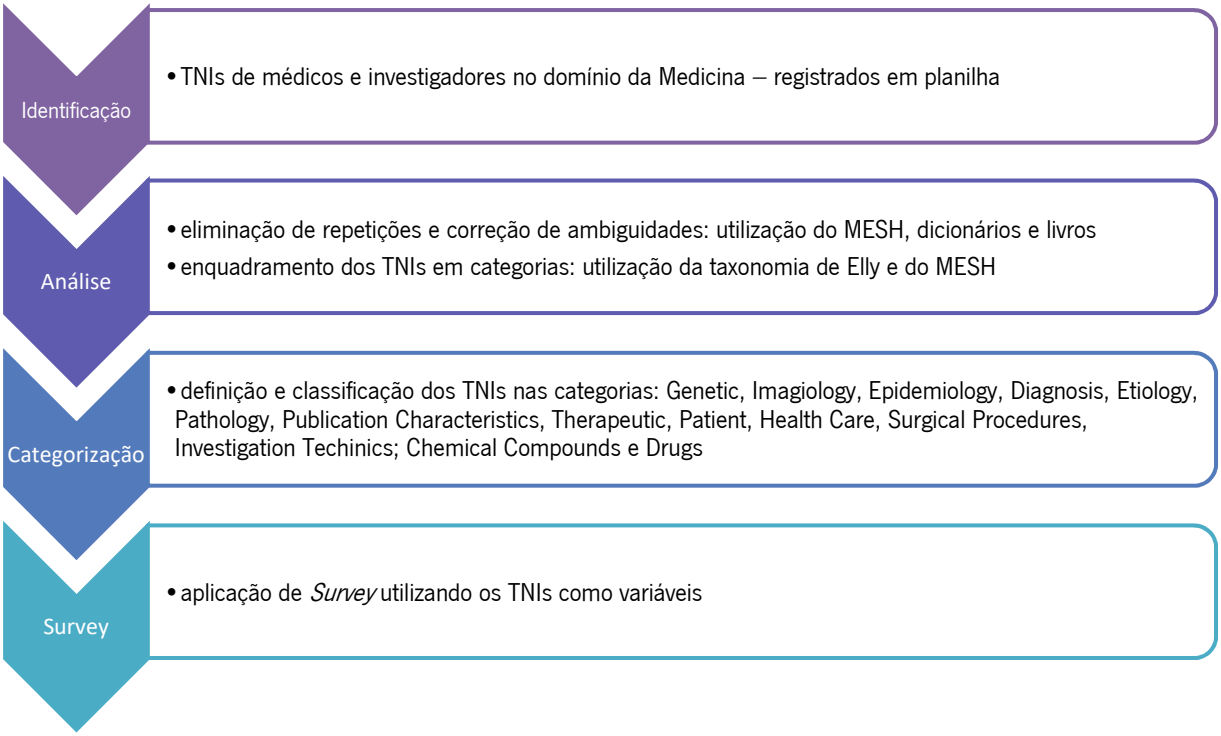


Figura 45. Atividades desenvolvidas para identificação dos tipos de necessidades de informação no domínio da Oncologia

Seleção dos TNIs

Objetivo: selecionar os TNIs a serem utilizados no Onco-MAP

Atividade:

a) analisar os resultados dos Estudos I, II e III

b) registrar e comparar os resultados em planilha;

c) identificar elementos de metadados adicionais necessários para gestão interna dos recursos nos RDCs. Essa tarefa está associada às atividades de gerência dos RDCs desenvolvidas por profissionais da informação (ex.: bibliotecários). A autora possui competência em relação a essa tarefa, não sendo necessário incluir outros atores para a sua realização[34]; e

d) eliminar ambiguidades e redundâncias.

Para que os TNIs resultantes dos Estudos I, II e III pudessem compor o PAM proposto por este estudo, denominado de Onco-MAP (Figura 46), eles foram analisados e categorizados utilizando como parâmetro, o DC Terms e a taxonomia dos TNIs em Oncologia (Figura 42).

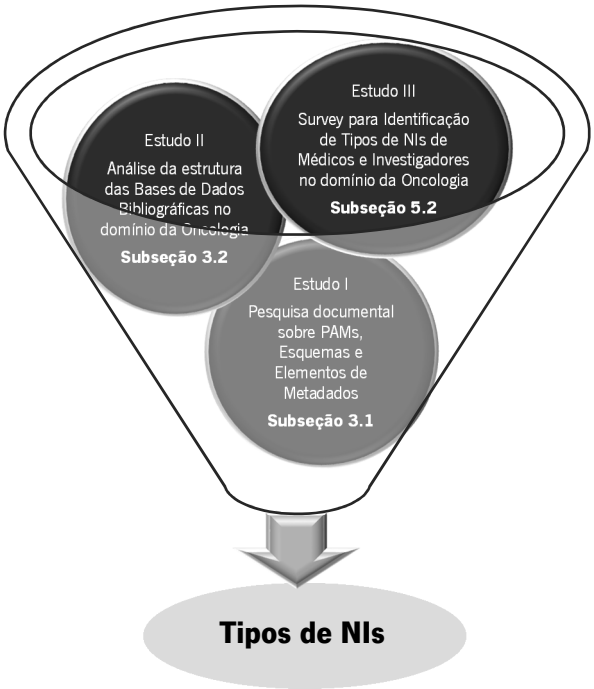


Figura 46. Seleção dos tipos de necessidades de informações.

Os resultados desses estudos estão reunidos na Tabela 71.

3434 Bibliotecária com 30 anos de experiências em normalização bibliográfica, pesquisa em bases em bases de dados (multidisciplinar e odontológica), atendimento no setor de referência e indexação de periódicos. Coordenadora e responsavábel pela implantação do repositório institucional da instituição a qual está vinculada.

Tabela 71. **Compilação dos resultados dos estudos desenvolvidos no âmbito da tese**.

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
Title	dc:title	dc: title	Title	Título do artigo
		General -> title	Journal title	Título da publicação
		titleInfo -> title	Journal abbreviation	
			Original title	
			Publication title	
Alternative Title	dcterms:alternative	dc:Alternative	Trasiliterated title	
Creator	dc:creator	dc: creator		Autor
		name -> namePart		
Contributor	dc:contributor	dc: contributor	Record owner	
Publisher	dc:publisher	dc: Publisher	Publisher	
		originInfo -> Publisher		
Date	dc:date	dc: date	Date Create	
	dcterms:created		Date Created	
	dcterms:modified		Date Modification	
	dcterms:dateSubmitted		Date Publication	
	dcterms:issued	originInfo -> dateIssued	Revision date	
	dcterms:dateAccepted		Version date	
Identifier	dc:identifier	dc: identifier	Unique identifier	
		location -> url	Unique identifier of investigator	
		identifier	Author name ID	
Language	dc:language	dc: language	Language	
		General ->language	Publication language	
Format/medium	dc:format	dc: format		

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
	dcterms:hasFormat			
Size	dcterms:Extent			
Availability	dcterms:available			
Subject	dc:subject	dc: subject	Subject	Assunto
		General -> keyword	MeSH Subject Heading	
		Classification -> keyword	Exploded sub-heading	
		subject -> topic	Topic Text Word	
			Geographic terms	
Coverage	dc:coverage	dc: coverage	Place name	
	dcterms:spatial	originInfo -> place		
	dcterms:temporal			
Description	dc:description	dc: description	Volume	File size
	dcterms:abstract	Educational -> description	Duration	
	dcterms:audience	dc:audience	Edition	
	dcterms:tableOfContents	mods:abstract	Page	
		originInfo -> edition	Pagination	
		Educational -> typical learning time	Start Page	
		General -> identifier -> catalog		
		General -> description		
		Educational -> typical age range		
		Educational -> typical age range		
Type	dc:type	dc: type	Publication type	
		Educational -> learning resource type	Conference	
			Media type	
Rights Description	uketdterms:embargotype			

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
Terms of use/access rights	dc:rights	dc: rights		
Relation	dc:relation	dc: relation		
	dcterms:hasPart			
	dcterms:hasVersion			
	dcterms:isPartOf	dc: isPartOf		
	dcterms:isReferencedBy			
	dcterms:isVersionOf			
	dcterms:references			
	dpla:replaces			
Bibliographic citation	dcterms:bibliographicCitation	dc:bibliographicCitation	Citation	Citações bibliográficas
			References	
			Cited Author	
			Citation analysis	
			Bibliographic citation	
Source	dc:source	dc: source	Source Title	
			Secondary source	
			Source Type	
Affiliation	vcard:org		Institution	
Funding agency			Nome do subsídio	
			Funding agency	
Notes			Reader comments	
			Notas	
			Comments	
Caption				Tabelas
				Figuras

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
Abreviaturas/siglas				Abbreviation
Factor de impacto				Factor de impacto
Descrição conteúdo				
Epidemiology				Causation/ association (etiology)
				Course/ prognosis
				Etiology
				Incidência por tipo de câncer
				Population statistics
				Prevalence/ incidence
				Prognosis
				Risk factors/ disease agentes
Genetics			Gene symbol Word	Anatomic location of target detected
			Gene symbol	Biological function (protein expression)
				Biological process and species
				Cell compartment and protein structure
				Cell component
				Color depth
				Comparison (tumor v adjacent normal)
				Expression level (protein expression)
				Gene id
				Gene id (DNA)
				Genetics
				Tissue type and disease
Pathology			Rare disease supplementary	Biologic mechanism
				Cause and risk factors
				Class of disease (cancer) (nature of disease)
				Date of staging (extent of disease)
				Diagnosis/etiology
				Disease prognosis

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
				Documentation for this cancer) (nature of disease)
				Specific findings TNM (stage) (extent of disease)
				Staging system (stage) (extent of disease)
				Status at staging (extent of disease)
				TNM stage (extent of disease)
				What previous diagnosis have been established
Surgical Procedure				What was the procedure carried out
Therapeutic				Administration in face of adverse effects
				Adverse effects
				Cancer directed/non cancer directed
				Contraindications
				Current therapeutics
				Dosage
				Efficacy/ indications/ drug of choice
				How to do it (treatment)
				Modification,additions/omission to therapy
				Protocol (therapeutic procedures)
Technique investigative				Crisp
				Safety/ contraindications (includes pregnancy and breast feeding)
Imagiology			Visualização de imagem	Anatomic frame of reference
			Image	Equipment
			Tipo de visualização imagem	Magnification

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
				Physical description
				Primary magnification
				Radioiodine
				Size of pixel
				Software
				What is the location and characteristics of the present lesion
Health Care			Consultas clínicas	Clinical trial information
				Clinically condition (orientation)
				Condition/ finding (management)
				Doctor-patient communication
				How to advise (management)
				How to approach difficult issue (doctor-patient communication)
				Management (not specifying diagnostic or therapeutic) (management)
				Patient compliance (doctor-patient communication)
				Practices of other providers (management)
				Community services (management)
Diagnosis				Anatomic region (specimen)
				Biomarker
				Body part (anatomy) on physical exam or imaging study
				Cause/ interpretation of clinical finding
				Criteria/ manifestations
				Description of processing
				Diagnosis, multiple
				Diagnostic criteria
				Pathology findings (specimen)

Atributo	Identificação de PAMs, esquemas e elementos de metadados		Identificação dos pontos de acesso de bases de dados bibliográficas (BDB)	Survey sobre tipos de NIs
	Pesquisa documental	Survey*		
				Tumor grade, size and stage at DX and resection
				Diagnosis, using laboratory
				Disease, biological mechanism
Etiology				Radiation
Patient			Age	Cancer status at death
			Gender	Family history of disease (family history of disease)
			Human or animal	How a patient should be managed disease information
				How documented (biopsy, cytology, radiology exam)
				Metastases (multiple)
				Method used to determine first recurrence
				Primary cause of death
				Site of metastasis
				Type of first recurrence (local, distant, biochemical)
				Vital status
Chemical composity and drugs			Pharmacological Action	Biological effect and mechanism of action
			Compostos químicos e drogas	Adverse effects
			Brand name	
			Registry number/ name of subs	
			Chemical name	

6.2 TIPOS DE NECESSIDADES DE INFORMAÇÃO: IDENTIFICAÇÃO DE CLASSES E ATRIBUTOS

A categorização dos elementos de metadados (Estudo I) e dos pontos de acesso (Estudo II) em classes e atributos não requereu atividades adicionais além da identificação da finalidade e sua associação. No entanto, em relação aos TNIs dos médicos e investigadores, o procedimento foi diferenciado por requerer a clarificação e a categorização dos termos no domínio da Oncologia.

Devido a dificuldade em identificar as definições desses TNIs nos artigos de onde foram extraídos, buscas foram realizadas em sistemas de organização do conhecimento (NCI Thesaurus e SNOMED-CT) e em livros para clarificar o conceito dos termos.

Para a classificação dos termos em classes e atributos, e a definição das relações, considerou-se a taxonomia dos TNIs (Figura 42) e os vocabulários controlados NCI Thesaurus e SNOMED-CT. Com este procedimento também foi possível clarificar e categorizar esses termos no domínio da Oncologia, visto que houve dificuldade em identificar propriedades em esquemas de metadados que representassem a maior parte dos TNIs.

Outra característica observada na busca pelos TNIs que refletiu na categorização foi a apresentação de questões formuladas por médicos e investigadores nos artigos científicos. Nesses casos, foi necessário traduzir as questões para termos. Em outros casos, foi identificado que a semântica ou a sintaxe dos TNIs poderia ser mais bem representada se substituída por termo do NCI Thesaurus, não deixando de considerar a taxonomia dos TNIs. Exemplo: uso do termo terapêutica e não tratamento. Esse processo está explicitado na árvore de decisão para sistematização de conceitos (Figura 47).

Motivação para adoção das ferramentas utilizadas nesta fase do estudo:

a) NCI Thesaurus – O Thesaurus do NCI é uma terminologia de referência e uma ontologia biomédica usada por um número crescente de institutos de Oncologia e outros sistemas. Abrange vocabulários para cuidados clínicos, translacional e investigação básica, informação pública e atividades administrativas. O formato adotado é OWL. O Thesaurus do NCI fornece definições, sinônimos e outras informações sobre quase 10.000 cânceres e doenças relacionadas, 8.000 agentes únicos e terapias de combinação e uma ampla gama de outros tópicos associados ao câncer e à investigação biomédica.

   O NCI Thesaurus inclui terminologias obtidas de outras fontes, como GO-Gene Ontóloga, HL7 Vocabulary, ICD-10 International Classification of Diseases, SNOMED e UMLS, entre outras. É dividido nas seguintes hierarquias: *Abnormal Cell*; *Activity*; *Anatomic Structure*; *System*, *Or Substance*;

Biomedical Pathway; *Biological Process*; *Chemotherapy Regimen Or Agente Combination*; *Conceptual Entity*; *Diagnostico Prognostic Fator*; *Disease*, *Disorder Or Finding*; *Drug*, *Food*, *Chemical or Biomedical Material*; *Experimental Organism Anatomical Concept*; *Experimental Organism Diagnosis*; *Gene*; *Gene Product*; *Manufactured Object*; *Molecular Abnormality*; *NCI Administrative Concept*; *Organism*; *Property or Attribute*; *e Retired Concept.*

b) Nomenclatura Sistematizada de Medicina – Termos Clínicos (SNOMED) – é mantida pela Organização Internacional para o Desenvolvimento de Normas de Terminologia em Saúde. Um dos seus objetivos é suportar a interoperabilidade semântica entre sistemas a partir da codificação de termos clínicos e mapeamento de terminologias. Possui as seguintes características: maior abrangência para a terminologia clínica multilíngue com foco no prontuário eletrônico de pacientes e assistência ao paciente; tomada de decisão e comunicação no contexto da saúde (Galvão & Ricarte, 2013).

   A SNOMED é dividida nas seguintes hierarquias: *Clinical Finding*; *Procedure*; *Observable Entity*; *Organism*; *Substance*; *Pharmaceutical/Biologic Product*; *Specimen*; *Special Concept*; *Physical Object*; *Physical Object*; *Event*; *Environments and Geographical Locations*; *Social Context*; *Situation with explicit context*; *Staging and Scales*; *Qualifier Value*; *Physical Force*; *Attribute* (https://www.nlm.nih.gov/healthit/snomedct/index.html).

O resultado obtido após a sistematização dos conceitos é apresentado na Tabela 72, em que constam as classes, os atributos, a conceitualização e o identificador que aponta para a fonte que proporcionou aquele termo. No total foram identificados 185 termos, incluindo classes e atributos.

Vale salientar que o objetivo principal destas atividades não foi oferecer uma nova definição para os termos, mas alinhar os termos ao domínio da Oncologia, buscando preservar a associação do termo às categorias apresentadas no questionário submetido aos entrevistados (Diagnóstico, Tratamento, Imagiologia, etc.), tendo em vista a estrutura de um artigo científico.

Quando o termo não era identificado no NCI Thesaurus, utilizava-se o SNOMED-CT. No caso de o termo indicado pelos entrevistados não ter sido localizado em nenhuma das duas ferramentas, este foi mantido e a conceitualização foi realizada com base em outras fontes: glossário temático, controle de câncer (Brasil. Ministério da Saúde, 2013), dicionário médico (Stedman, 2003), vocabulário MESH, artigos e livros (Sobrinho-Simões, 2016). A opção por este procedimento foi de ser coerente com a proposta do estudo: identificar os TNIs dos médicos e investigadores, portanto não justificando a eliminação do termo por ele não ser identificado em um vocabulário controlado.

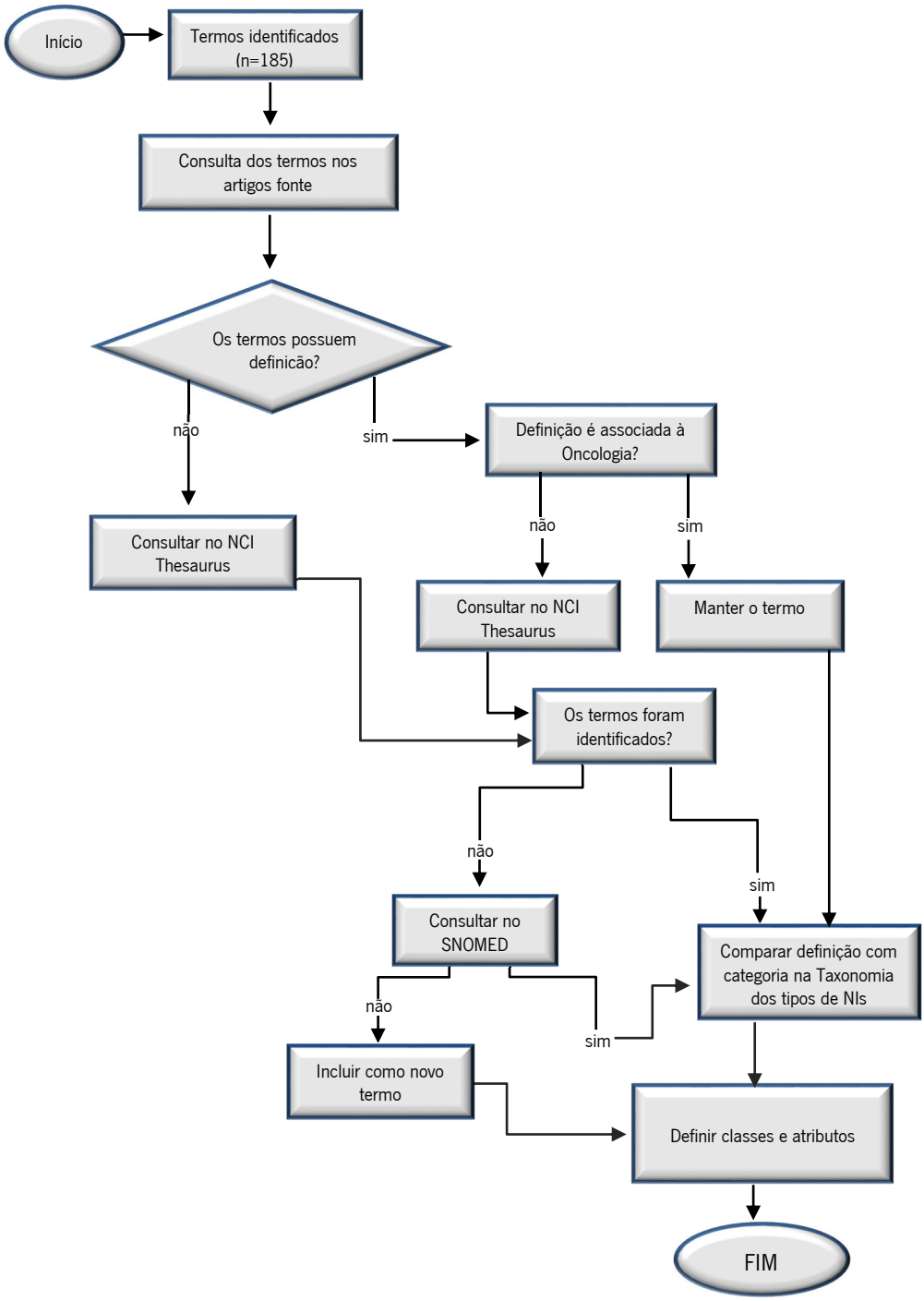


Figura 47. Processo de sistematização dos conceitos.

Tabela 72. **Tipos de necessidades informacionais: definição de classes e atributos**

Classe/Atributo		Descrição	Id
Classe	**Publisher**	Publishing company	4
Atributo	Name	Name of entity responsible for making the resource available.	3
	Homepage	A homepage for publisher.	1
	Location	The spatial topic of the resource.	1
	Date created	Date of formal issuance of the resource.	1
Classe	**Journal**	A periodical of scholarly jornal	1
Atributo	Title	A name given to the journal.	1
	Abbreviation	The abbreviated form representing a resource.	2
	Identifier	An unambiguous reference to the resource.	1
	Publisher	An entity responsible for making the resource available.	1
	Date Created	Date of formal issuance of the resource.	1
	Homepage	A homepage for resource.	2
	Score (metric)	Represents the metric for resource.	2
	Type	Publishing company	1
Classe	**Scientific Article**	A scholarly academic article, typically published in a journal	4
Atributo	Title	A name given to the resource.	1
	Identifier	An unambiguous reference to the resource within a given contexto	1
	Volume	Volume of a journal.	2
	Issue	Issue of a journal.	2
	Page range	Page range of a document.	2
	Date Issued	Date of publication of the resource.	1
	Author	An entity primarily responsible for making the Resource	1
	Abstract	A summary of the resource.	1
	Subject	The topic of the resource.	1
	Language	A language of the resource.	1
	References	A related resource that is referenced, cited, or otherwise pointed to by the described resource.	1
	Appendix	Text or document prepared by the author	4
	Illustration	A image, graphic or table within a journal.	2
	Type	The nature or genre of the resource.	1
	Format	The file format, physical medium, or dimensions of the resource.	1
	Rights	Information about rights held in and over the resource.	1
	Access Mode	Declares the degree of openness of a resource.	4
	Score (metric)	Represents the metric being observed.	2
	Grant Number	A data property specifying the grant number of a grant provided by a funding agency	2
	Level of Evidence	Strength of evidence of the data used to formulate the guideline.	4
	Conflit of Interest	A situation in which an individual might benefit personally from oficial or professional actions	4
	Relation	A related Resource	1
	Status	Status of a Resource	4
	File size	File size corresponding to the file of article	1
	Duration	Duration of the audio or video corresponding to the file of article	1
Classe	**Author**	A resource that acts or has the power to act	1
Atributo	Name	An entity primarily responsible for making the resource.	2
	Identifier	An unambiguous reference to the resource within a given context.	1
	Affiliation	An organization that this person is affiliated	1
	Mailbox	A personal mailbox.	4
	Score (metric)	A numeric values measuring performance or quality of something.	2

Classe/Atributo		Descrição	Id
Classe	**Institution**	An organization, establishment or foundation.	
Atributo	Name	Name of Institution.	2
	Homepage	A homepage for organization.	1
	Location	Physical location of the item.	1
	Type	The nature or genre of the resource.	1
Classe	**Funding agency**	Insitutions or organizations that support scientific or technologic research	1
Atributo	Name	Name of agency	1
	Homepage	A homepage for Organization	1
	Location	Physical location of the institution.	1
	Type	The nature or genre of the resource.	1
Classe	**Design Research**	Describes how the study performed to answer the research	1
Atributo	Data collection technique	Referes to process of gathering data for testing or research purposes.	4
	Research method	A method or procedures used in a research.	2
	Statistical technique	A statistical technique used in a research.	4
	Geographic location	Place where the study was developed.	1
	Sample	The sample of study.	3
	Sample size	The sample size.	3
	Study Subject	A subject investigated by the study.	3
Classe	**Illustration**	A numbered and captioned image, graphic or table within a journal.	3
Atributo	Caption	The caption for this object.	1
	Description	An account of the resource.	1
	Type	The nature or genre of the resource	1
Classe	**Note**	A note about a Resource	1
Atributo	Author	Na entity primarily responsible for making the resource.	1
	Date created	Date of creation of the resource	2
	Mailbox	A personal mailbox	1
	Relation	A related resource.	1
	Description	An account of the Resource	1
Classe	**Gene**	The functional and physical unit of heredity passed from parente to offspring.	3
Atributo	Expression level	The determination of the level of gene transcription in a biological specimen.	3
	Gene id	A unique gene name within a specific repository, database, or collection.	3
	Gene symbol	A unique gene name approved by an organism specific nomenclature committee.	2
Classe	**Organism**	A living entity.	3
Atributo	Age	Elapsed time since birth.	2
	Ethnicity	Members share a common history and a sense of identification with the group	2
	Gender	The gender of this agent.	2
	Race	An arbitrary classification of a taxonomic group that is a division of a species.	2
	Description	Description of a living entity.	1
Classe	**Anatomy**	Physical structure of animals and plants	3
Atributo	Anatomic region	Identifies areas of the body.	3
	Biological process	An activity occurring within an organism, between organisms or among organisms and the mechanisms underlying such events.	3
	Cell component	The components of the cell.	3
	Protein structure	The amino acids and their 3-D arrangement in constituting a protein.	3
	Tissue type	Describes collections of differentiated cells, such as epithelium; connective tissue; muscles; nerve tissue.	3
	Description	Describes the referenced anatomy.	1

Classe/Atributo		Descrição	Id
Classe	**Neoplasm**	A benigno or malignant tissue growth resulting from uncontrolled cell proliferation	3
Atributo	Cause	One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence.	3
	Classification	The classification of cancers by histological type; primary site; location where it was first developed.	3
	Date of staging	Date of he process of defining at what point in the natural history of a malignant disease a patient is when a diagnosis is made.	3
	Grade	A position on a scale of intensity or amount or quality.	3
	Lesion characteristic	Characteristic of localized pathological or traumatic structural change, damage, deformity, or discontinuity of tissue, organ, or body part.	3
	Biologic mechanism	Biological mechanism associated with disease.	3
	Metastasis	The spread or migration of cancer cells from one part of the body to another.	3
	Recurrence	If the cancer returned.	3
	Site	A position in which cancer develops (histologic and morphologic).	3
	Size	The physical magnitude of tumor.	3
	Status at staging	Level of extent of the cancer within the body.	3
	Type of cancer	Type of cancer.	3
Classe	**Disease**	A health problem with certain characteristics or symptoms	3
Atributo	Cause	One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence.	3
	Characteristic	The distinguishing of a disease aspect.	3
	Class of disease	Classification of diseases	3
	Complication	Any disease or disorder that occurs during the course of, or because of, another disease, treatment, or procedure.	3
	Description	Description of the disease.	1
	Disease agent	Any vírus, micro-organism, or other substance causing disease.	3
	Biological mechanism	Biological mechanism associated with disease.	3
	Prognosis	Prognosis.	3
	Rare disease supplementary	Rare disease identification.	2
	Risk factors	A modifiable or non-modifiable fator that increases the risk of a person contracting a condition.	3
Classe	**Diagnosis**	The investigation, analysis and recognition of the presence and nature of disease, condition, or injury from expressed signs and symptoms.	3
Atributo	Category	A category for the item.	3
	Clinically condition	Specifying the health condition(s) of a patient, medical study, or other target audience.	3
	Condition/ finding	Clinical, laboratory or molecular evidence, or lack of evidence of the disease.	3
	Interpretation of clinical finding	Interpretation of clinical finding.	3
	Criteria/ manifestations	Criteria based on patient symptoms and signs, test results.	3
	Pathology findings	Pathology findings.	3
	Previous diagnosis	Previous diagnosis.	3
	Sign or symptom	A sign o r symptom of this condition.	3
	Vital status	The state or condition of being living or deceased; also includes the case where the vital status is unknown.	3

Classe/Atributo		Descrição	Id
Classe	**Diagnostic procedure**	A medical procedure intented primarily for diagnostic, as opposed to therapeutic, purposes.	3
Atributo	Biomarker	Biomarker used.	3
	Chemical agent administration	Administration of chemical agent in exams.	3
	Diagnostic criteria	A standard from which a judgment concerning a medical diagnosis can be established.	3
	Staging System	A system adoped to describe the extent of cancer in a patient's body. Ex.: TNM Staging System.	3
	Technique	Encompass all investigations and tests intended to identify the cause of disease. Ex.: laboratory tests and imaging techniques, Differential diagnosis.	3
	Test	Process of determining the disease state of a person. Ex.: cancer mutation test, blood chemistry test, etc.	3
	Type of procedure	The type of procedure, for example surgical, noninvasive or percutaneous. Ex.: Clinical diagnosis; computer-aided diagnosis.	
Classe	**Imagiology**	Any medical imaging modality typically used for diagnostic purposes.	3
Atributo	Anatomic frame of reference	Image acquisition module.	3
	Caption	The caption for this object.	1
	Color depth	Positive integer indicating the desired minimum total number of bits per color channel used to present a pixel.	3
	Date created	Date of creation of the resource.	3
	Description	An account of the image.	3
	Equipment	Equipment used to analyze, enhance, interpret or display digital images	3
	Format	The duration of the item.	1
	Magnification	Magnification is related to scaling up visuals or images to be able to see more detail.	3
	Physical description	Physical description of the image.	3
	Primary magnification	Initial magnification.	3
	Processing	The use of data processing systems to analyze, enhance, interpret or display digital images.	3
	Provenance	Ownership and custody of the resource of since its creation.	3
	Size of pixel	Image size expressed in pixels.	3
	Software	A computer program in source of compiled form.	3
	Type	The nature or genre of the resource.	3
	Visualization Mode	Options for visualization	2
Classe	**Therapeutic**	Relating to or involved in therapy	3
Atributo	Adverse effect	An unexpected medical problem that happens during treatment with a drug or other therapy.	3
	Contraindication	A contraindication for this therapy	3
	Current therapeutic	The name of the therapy that is currently being administered.	3
	Drug of choice	The class of drug this belongs to (ex., statins)	3
	Efficacy	The ability of a chemical agent or a device to treat a disease.	3
	Indication	A fator that indicates use of this therapy for treatment and/or prevention of a condition, symptom, etc.	3
	Monitoring	How a person or a condition is assited or verified.	3
	Patient compliance	The act of following a medical regimen or schedule correctly and consistently, including taking medicines or following a diet.	3
	Prevention	Describes methods of preventing cancer including limit of exposure to carcinogens and infectious agents, diet, education, etc	3

Classe/Atributo		Descrição	Id
	Procedure	Describes a treatment or action taken to prevent or treat disease	3
	Protocol	Describes a rule which guides how an activity should be performed.	3
	Type of treatment	Cancer treatment adopted to non directed or directed cancer Ex.: control, remove or destroy cancer or prolong a patient's live, alleviate pain.	3
Classe	**Phamacological substance**	Any substance, other than food, that is used to prevent, diagnose, treat or relieve symptoms of a disease or abnormal condition	2
Atributo	Adverse effects	An unexpected medical problem that happens during treatment with a drug or other therapy.	3
	Drug	Specifying a drug or medicine used in a medication procedure	3
	Biological effects	Biological effect that the any product provide.	3
	Brand name	Brand name.	2
	Chemical name	Chemical name.	2
	Dosage	The value of the dose.	3
	Mechanism of action	The specific biochemical interaction through which this drug or supplement produces its pharmacological effect	3
	Pharmacalogical action	Actions of drugs and their effects on living organisms	3
	Radioiodine	Use of radioiodine.	3
	Reaction	A description of the action in response to some kind of stimulus.	3
Classe	**Epidemiology**	The study of the patterns, causes, and control of disease in groups of people	3
Atributo	Causality/ association	Relationship exists between a cause or effect.	3
	Clinical trial information	Report of a pre-planned medical or veterinary clinical case study involving a randomized, evidence-controlled clinical trial.	3
	Etiology	The cause of a disease or abnormal condition.	3
	Genetic	Interdisciplinary clinical, epidemiologic, genetic, and laboratory studies of persons, families, and populations at high risk of cancer.	3
	Incidence by type of cancer	The relative frequency of occurrence of cancer.	3
	Population statistics	A value, which represents a particular population characteristic.	3
	Prevalence	The total number of cases of a given disease in a specified population at a designated time.	3
	Primary cause of death	The first significant event which ultimately led to death.	3

Nota: Identificação (ID) da fonte de onde o TNI foi extraído: (1) PAMs, (2) pontos de acesso de BDBs, (3) *Survey* com médicos e investigadores e (4) contribuição da autora.

Esta etapa do trabalho corresponde à identificação de requisitos (funcionais e informacionais) que possibilitam o desenvolvimento do artefacto – o PAM (fase Ambiente do DSR). As próximas Seções referem-se ao desenvolvimento do PAM e a sua validação (fase do Desenvolvimento do Artefacto). A Figura 15, que representa o enquadramento metodológico adotado para o desenvolvimento deste estudo, é reinserida aqui com o intuito de demonstrar a interação entre as atividades desenvolvidas na elaboração desta tese (Figura 48), em que cada fase do DSR é relacionada à Seção correspondente, de modo a explicitar a interação entre elas.

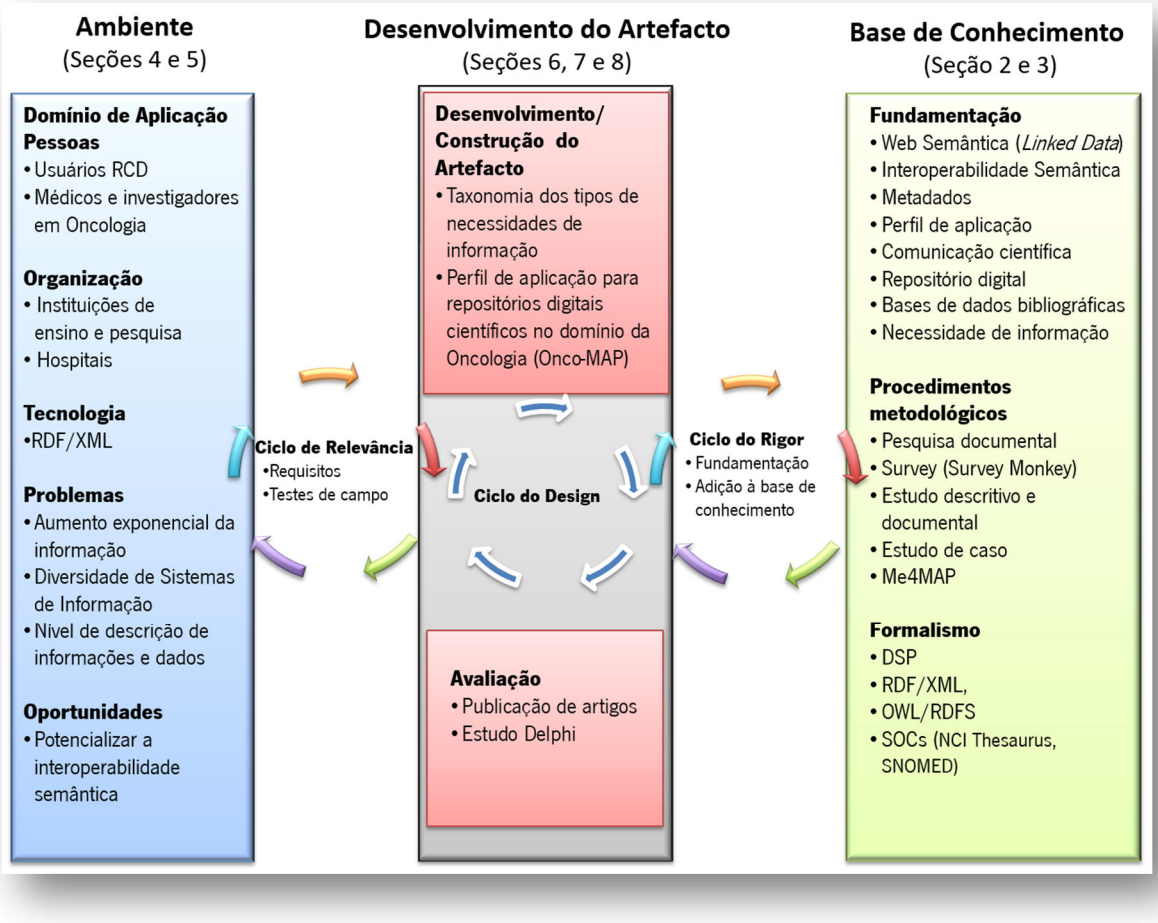


Figura 48. As fases do DSR e a tese

SÍNTESE DA SECÇÃO 6

Reúne os resultados obtidos com os três estudos desenvolvidos no âmbito desta tese e apresenta os procedimentos adotados para a tradução dos TNIs em classes e atributos a serem incorporados ao perfil de aplicação proposto por esta investigação. A seguir será apresentada a Seção 7, correspondente à fase de Desenvolvimento do Artefacto do DSR, ou seja, a elaboração do PAM a partir dos TNIs (Tabela 72).

7 DESENVOLVIMENTO DE UM PERFIL DE APLICAÇÃO NO DOMÍNIO DA ONCOLOGIA

> Para a comunicação em geral, é importante que a pessoa B seja capaz de compreender o que a pessoa A está a dizer. Para este entendimento mútuo, é necessária uma base comum, um léxico básico com o conhecimento do significado das coisas. A partir desse ponto, pode-se começar a raciocinar. Para apoiar a comunicação científica com a utilização de repositórios, os repositórios deverão falar a mesma língua e é, portanto, essencial criar uma base comum (Digital Repository Infrastructure Vision for European Research - DRIVER Project, 2008, p. 2)

Para o desenvolvimento do PAM proposto para este estudo – Onco-MAP –, foi utilizado o Método Me4MAP, que segue os princípios do Enquadramento Singapura adotados pela DCMI para desenvolvimento de PAMs. O enquadramento Singapura "[...] *is a framework for designing metadata applications for maximum interoperability and for documenting such applications for maximum reusability*" (Nilsson, Baker & Johnston, 2008, p.1). Esse enquadramento apresenta a descrição de um conjunto de componentes necessários para o desenvolvimento de perfil de aplicação, assim como normas relacionadas com o modelo de domínio e os fundamentos da Web Semântica.

O Me4MAP busca estabelecer procedimentos e sugerir técnicas para o desenvolvimento de um PAM (Malta & Baptista, 2013, Malta, 2014). O método apresenta etapas recomendadas pelo Enquadramento Singapura, com adição de orientações e abordagens que facilitam o entendimento e a elaboração de um perfil (Figura 49).

A elaboração deste perfil visa a:

a) potencializar a interoperabilidade semântica entre RDCs no domínio Oncologia;
b) benefíciar a comunidade de repositórios digitais (interoperabilidade entre os RDCs e os demais sistemas de informação);
c) facilitar o processo de busca de artigos científicos, com utilização de vocabulários para domínios específicos (Ontologia) (NI);
d) favorecer o intercâmbio de informações entre os diversos profissionais que atuam no domínio da Oncologia:
   - integrar e compartilhar informações médicas altamente heterogêneas e conhecimento científico, utilizando as melhores práticas da Web Semântica (*Linked Open Data*); e
   - apoiar a investigação translacional e a Medicina Baseada em Evidências no atendimento ao paciente.

Para cada etapa e subetapa apresentada no Me4MAP, são incluídas as seguintes informações:

a) definição;
b) sequência da atividade;
c) denominação da atividade;
d) do que se trata;
e) como deve ser desenvolvido;
f) por quem;
g) produto entregue;
h) pré-requisitos para o desenvolvimento da atividade; e
i) resultados alcançados.

Dessa forma, o método permite que os desenvolvedores possam realizar adaptações para diferentes situações. Neste estudo, nem todas as etapas ou atividades são adotadas devido a algumas particularidades, como a existência de apenas um desenvolvedor. Assim, explicaremos as etapas do fluxo à medida que elas forem utilizadas.

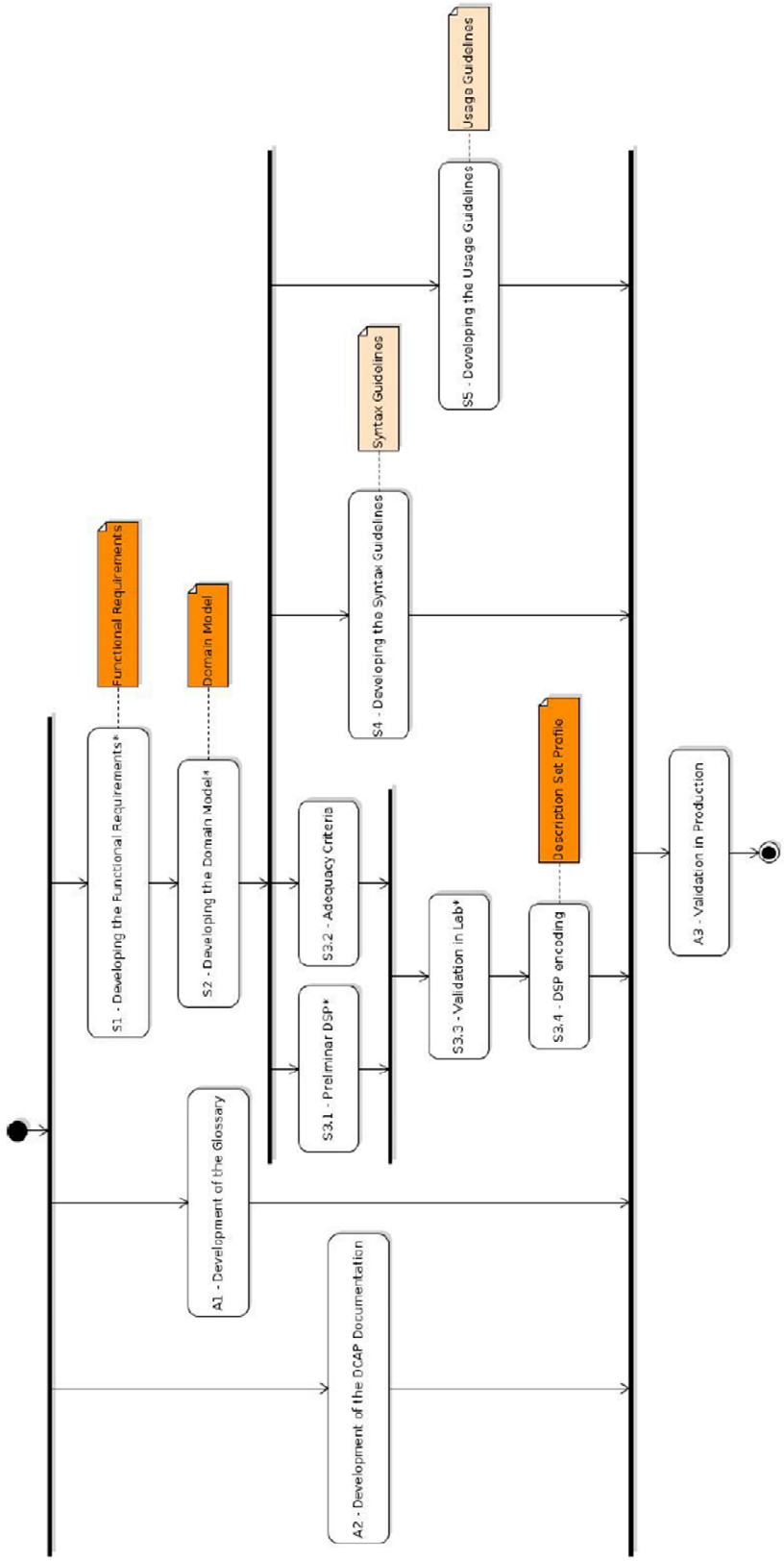


Figura 49. Esquema lógico para a construção do perfil de aplicação. Recuperado de Contributo metodológico para o desenvolvimento de perfis de aplicação no contexto da Web Semântica, de M. C. Malta, 2014, Universidade do Minho.

7.1 IDENTIFICAÇÃO DOS REQUISITOS

A primeira fase do desenvolvimento do PAM corresponde à definição dos requisitos e prevê a elaboração dos seguintes artefactos: documento visão; plano de trabalho; mapeamento do ambiente; requisitos de alto nível e modelo de casos de uso.

Para Leite (1994, p. 50-51), a identificação, ou elicitação de requisitos[35], pode ser conceituada como o fato de "[...] descobrir, tornar explícito, obter o máximo de informações para o conhecimento do objeto em questão" e está relacionada com as tarefas de "[...] identificação e fontes de informação, coleta de fatos e comunicação". Quanto à identificação de fontes de informação, o autor salienta que os usuários não são as únicas fontes de informação. Merecem também atenção os documentos relativos ao sistema, livros sobre o tema e outros sistemas.

A partir da identificação das fontes de informação, dar-se-á a coleta de fatos. Leite (1994) enumera, entre as técnicas possíveis: leitura de documentos, observação, entrevistas, questionários, análise de protocolo, participação dos atores, enfoque antropológico, reuniões, reutilização e recuperação do desenho de *software*. Dentre essas técnicas, destacam-se, no âmbito deste estudo, a leitura de documentos, questionário e a reutilização de resultados de outras investigações.

Em relação à técnica da reutilização, Leite (1994, p. 53) assevera que é "[...] extremamente produtivo se, ao procurar definir requisitos para um determinado *software*, pudéssemos reutilizar fatos já coletados [...]". Dessa forma, é possível reduzir o tempo gasto para a compreensão e modelagem, visto que os conceitos já se encontram disponíveis. Uma das vantagens defendidas pelo autor é contar com a reutilização de elementos já validados.

Nesses termos, este estudo atende pelo menos a três das técnicas elencadas por Leite (1994) para elicitação de requisitos. São elas:

a) fonte de informação: análise e obtenção de informação por meio de pesquisa documental;
b) questionário: foram aplicados dois *surveys*, um para médicos e investigadores no domínio da Oncologia e outro para gestores de repositórios institucionais; e
c) reúso: foram reutilizados resultados de estudos sobre TNI no domínio da Medicina e da Oncologia.

Dando seguimento ao método Me4MAP, a primeira etapa a ser elaborada é o documento visão.

---

[35] O autor utiliza o termo requisito de forma genérica. Entretanto, neste estudo, considera-se que os requisitos podem estar associados às funcionalidades do sistema (requisitos funcionais, não aprofundados pelo estudo) e às necessidades dos utilizadores (requisitos informacionais, foco deste estudo).

Documento visão

Os sistemas de informação, como bases de dados bibliográficas e repositórios digitais, já dispõem de ferramentas que podem prover os conteúdos necessários para o aperfeiçoamento e a atualização de muitos profissionais. No entanto, esses sistemas, em sua maioria, raramente são desenvolvidos com base nos TNIs dos seus utilizadores, o que pode comprometer a qualidade dos resultados de busca (Subseções 4.1 e 4.2).

A Oncologia é um domínio de conhecimento em que o aumento exponencial da informação, a complexidade do domínio, a multidisciplinaridade, o detalhamento da informação (granularidade), a qualidade dos metadados e o grau de interoperabilidade semântica têm impactado os resultados de busca em sistemas de informação (Seção 5). Nesse sentido, é proposto o desenvolvimento de PAM, com base nos TNIs de médicos e investigadores em Oncologia.

Plano de trabalho

O Me4MAP, por possuir uma estrutura lógica, possibilita restringir as etapas para a elaboração de um PAM de acordo com a necessidade do desenvolvedor e das particularidades do projeto. Essa característica permitiu que este estudo seguisse apenas os procedimentos de determinadas etapas.

O PAM, artefacto a ser desenvolvido neste estudo, será aplicado ao domínio da Oncologia no contexto dos RDCs. Esse artefacto objetiva proporcionar qualidade à interoperabilidade semântica entre os RDCs e entre outros sistemas, de maneira que utilizadores no domínio da Oncologia possam acessar diferentes sistemas simultaneamente com maior eficácia. Este PAM inclui TNIs dos utilizadores como forma de contemplar especificidades inerentes ao domínio da Oncologia e de disciplinas correlatas.

Conforme abordado anteriormente, os tipos de recursos armazenados nos RDCs podem variar de instituição para instituição. No entanto, há uma preponderância de artigos científicos, *papers* apresentados em congressos, dissertações e teses (recentemente a inclusão de dados de investigação). Quanto à estrutura dos RDCs, a descrição desses recursos é realizada por meio de elementos de metadados, que podem ser usados de forma genérica, como o DC, ou voltados para objetos e ou objetivos específicos, como LOM, METS, MODS, FOAF e Premis, entre outros.

Quanto aos atores que atuam nos RDCs no domínio da Oncologia, podem ser professores, médicos, investigadores, técnicos e administradores dos RDCs (bibliotecários, informáticos), sendo que este estudo foca apenas nos investigadores e médicos (Seção 5).

Identificação de Requisitos

De acordo com a análise da literatura e do *Survey* com os administradores de RDCs (subseções 4.1 e 4.2), é possível afirmar que os RDCs devem (Andrade & Baptista, 2015): estar baseados na Web; ser interoperáveis; permitir o depósito de recursos; possibilitar a descrição dos recursos; relacionar seus dados com os de outros sistemas de informação; oportunizar, aos motores de busca, o acesso aos metadados; possibilitar a pesquisa dos recursos; permitir aos utilizadores a busca de recursos (por autor, título, assunto, tipos de investigação, evidências científicas, procedimentos médicos e tipos de amostras, entre outros); propiciar a leitura e o *download* dos recursos selecionados.

A Tabela 73 apresenta questões, estabelecidas por Coyle & Baker (2009, p. 1), respostas a essas questões e os requisitos funcionais correspondentes, os quais foram utilizados como base para o desenvolvimento do Onco-MAP.

Tabela 73. **Identificação dos requisitos funcionais**

Questões	Respostas às questões	Requisitos funcionais
O que você deseja realizar com a sua aplicação?	● potencializar a interoperabilidade semântica entre sistemas de informação	● ser interoperável
Quais são os limites da sua aplicação? O que não vai tentar fazer?	● repositórios digitais científicos. O limite será estabelecido pelo modelo de dados implementado no repositório ● não expandir para outros sistemas de informação, ex.: bibliotecas digitais, museus, arquivos	● limitar aos repositórios digitais científicos
Como você deseja que a aplicação atenda seus utilizadores?	● disponibilizar o acesso aos dados que descrevem os artigos no repositório ● dirimir possíveis problemas em relação à descoberta de artigos científicos proporcionando resultados com maior relevância	● possibilitar pesquisa artigos científicos ● acessar artigos científicos ● permitir o depósito de recursos
Seu aplicativo precisará executar ações específicas, como classificação ou download de dados em formatos específicos?	● download de arquivos no formato que foi armazenado nos repositórios já é uma funcionalidade intrínseca	● classificar artigos por relevância ● permitir download dos artigos em PDF ou ePUB
Quais são as principais características dos seus recursos, e como isso afeta sua seleção de elementos de dados?	● possuem autoria declarada ● inclui informações sobre o(s) autor(es), como o recebimento de algum tipo de financiamento e/ou vínculo empregatício ● apresenta uma estrutura que pode ser dividida em: título, resumo, introdução, método, resultados, conclusão, referências, apêndices e citações	● utilizar elementos de metadados para descrição de artigos científicos ● adotar normas para representação descritiva e temática do artigo - vocabulários e normas técnicas ● permitir aos utilizadores a busca de recursos por autor, título, assunto, tipos de investigação, evidências científicas, entre outros

Questões	Respostas às questões	Requisitos funcionais
	● é publicado por instituição e/ou empresa denominada de editora ● revisado por pares ● pode ter acesso restrito ou livre ● pode incluir tabelas, gráficos, figuras e notas ● pode trazer conteúdo técnico, teórico ou empírico ● utiliza pontos de acesso para descrição ● possui tipologia definida: artigo de revisão, investigação, relato de caso clínico, etc. ● apresenta resultados de investigações ou relatos clínicos ● versa sobre um ou mais tópico (diagnóstico, tratamento, genética, câncer, tipos de doenças, indivíduos, substâncias químicas (medicamentos), epidemiologia, imagiologia) ● faz uso de vocabulário técnico e especializado	
Quais são as principais características dos seus utilizadores? Eles estão associados a uma instituição particular ou estão atendendo um público em geral? Todos falam o mesmo idioma? Quão especialista são eles em relação aos dados que sua aplicação irá gerenciar? Quão especialista são sobre o tipo de recursos descritos?	● estão associados ao contexto acadêmico-científico e clínico ● não falam o mesmo idioma ● são profissionais que atuam no domínio da Oncologia ou em domínios relacionados ● em razão da formação acadêmica dos usuriaos é possivel deles terem utilizado ferramentas similares	● permitir livre acesso ao sistema ● possuir interface multilingue ● ser intuitivo e simples ● manter padrão similar ao repositorios existentes.
Existem padrões de comunidades que precisam ser considerados?	● Sim, os da comunidade médica e a dos investigadores, como o MESH, SNOMED e o NCI Thesaurus. Instrumentos utilizados para identificação de termos controlados e/ou definições terminológicas e hierárquia de conceitos	● Permitir o uso de vocabulários controlados

Os requisitos funcionais foram limitados ao que o repositório digital pode oferecer e suas características de acordo com a literatura consultada e com base no Survey com os RDCs (Andrade & Baptista, 2016).

Os requisitos informacionais, ou seja, a identificação dos TNIs que surgem a partir da contribuição dos especialistas do domínio da Oncologia, derivam dos requisitos funcionais. Segundo Victorino (2011, p. 158), "[...] é possível traduzir os requisitos de informações em especificações de um recurso para a representação da informação e do domínio". Neste caso, eles são fatores preponderantes para a

definição dos PAMs. A identificação desses dois tipos de requisitos contribui para uma visão holística do sistema, em que são averiguadas as funcionalidades e as necessidades de informação em relação ao conteúdo de um domínio específico (Seção 5).

7.2 DESENVOLVIMENTO DO MODELO DE DOMÍNIO

Esta etapa da elaboração do PAM engloba duas atividades, o desenvolvimento do *Environmental Scan*, e do modelo de domínio.

O *Environmental Scan*, ou "mapeamento do domínio", é um "relatório que contém uma revisão dos esquemas de metadados que estão disponíveis em qualquer serialização da Web Semântica (RDF/XML e Turtle) e que poderão servir às necessidades do modelo de domínio" (Malta & Baptista, 2017, p. 105).

O *Environmental Scan* foi realizado com base em registros de metadados, como o Language open Vocabulary (LOV), Metadata Registry e Schema.org. Foram avaliados esquemas e elementos de metadados utilizados por repositórios digitais científicos (Tabela 35) em consonância com os sistemas de organização do conhecimento NCI Thesaurus e o SNOMED, tendo como referência os termos médicos.

7.2.1 **Modelo de domínio**

O modelo de domínio, definido por Coyle e Baker (2009, p. 1) como sendo a "[...] caracterização dos tipos de coisas descritas por metadados e seus relacionamentos", foi desenvolvido com base nos três estudos realizados no decorrer desta tese:

a) Estudo I. Pesquisa documental sobre PAMs, Esquemas e Elementos de Metadados (Subseção 4.1);
b) Estudo II. Análise da estrutura das BDBs no domínio da Oncologia (Subseção 4.3); e
c) Estudo III. Survey para Identificação de Tipos de NIs de Médicos e Investigadores no domínio da Oncologia (Subseção 5.2). A compilação dos resultados desses estudos pode ser visualizada na Tabela 71.

O modelo de domínio foi desenvolvido para representar as informações contidas em artigos científicos no domínio da Oncologia no contexto dos repositórios digitais científicos. As informações foram representadas em um modelo de dados, em que as classes correspondem a um conjunto de elementos que possuem propriedades em comum e as propriedades se referem a aspectos que descrevem a classe.

Inicialmente, esses termos foram representados em classes e em atributos de um artigo científico em que foram consideradas características próprias do domínio da Oncologia e do contexto dos RDCs. Nesse sentido, ficou estabelecido que um artigo científico pode possuir classes ou atributos em relação:

a) à estrutura física: identificador (DOI); volume; número; intervalo de páginas; data de publicação do artigo; título; autoria; ilustrações; referências; e apêndices, que podem corresponder às imagens, aos questionários, entre outros;

b) à representação semântica do conteúdo: resumo; palavras-chave; tópicos de estudo, como: gene, câncer, diagnóstico, procedimentos diagnósticos, imagiologia, organismo, anatomia, patologia, terapêutica, substância farmacológica, epidemiologia; desenho de investigação;

c) às características do artigo: tipo de artigo (artigo original, revisão, comunicação e casos clínicos); métricas, nível de evidência científica; modo de acesso (livre, restrito ou embargado); direitos autorais; *status* (*preprint* ou *postprint*); indicação se foi produzido a partir de outro trabalho, como uma tese ou projeto de investigação; conflito de interesses; agência de financiamento associada ao artigo e o número do processo de financiamento;

d) às características do arquivo do artigo: duração para versão do artigo em áudio, formato e tamanho do arquivo.

Os termos representados como classes foram definidos a partir do relacionamento de propriedades e do tipo de informação que representam. A Web Semântica é orientada à propriedade. Nesse caso, as classes existem para restringir a aplicação das propriedades (*property centric*). "As classes surgem como domínios ou contradomínios de propriedades".[36]

Nesse sentido, dentre os termos relacionados nas alíneas a), b), c) e d), foram definidos como classes:

a) *Scientific Article*, que representa um artigo científico;

b) *Author* – esta classe possui propriedades e está associada à classe *Institution* e à classe *Scientific Article*. O autor pode possuir vínculo estudantil ou de trabalho ou pode ser uma pessoa jurídica.

c) *Design Research* – representa os procedimentos adotados para o desenvolvimento de um estudo. Pode estar relacionada no todo ou em parte com um artigo, como casos clínicos que não têm definição clara em relação a um método, à coleta de dados, mas possuem uma amostra que em geral se restringe a um único indivíduo;

[36] Informação verbal emitida por Ana Alice Baptista.

d) *Gene*, *Neoplasm*, *Diagnosis*, *Diagnosis Procedure*, *Imagiology*, *Organism*, *Anatomy*, *Disease*, *Therapeutic*, *Pharmacologic Substance*, *Epidemiology* – estas classes possuem propriedades; podem estar associadas ao *Scientific Article.* Por exemplo: um artigo pode ser representado pelas classes *Gene*, *Anatomy*, *DiagnosticProcedure* e *Neoplasm* (https://www.hindawi.com/journals/bmri/2017/2563085/);

e) *Funding Agency* – está associada à classe *Author* e à propriedade *grantNumber.* A propriedade *grantNumber* possui a particularidade de não ser uma propriedade natural da classe *Scientific Article.* Pode estar associada aos autores ou a um projeto/tese que originou o artigo científico, mas, com o objetivo de simplificar a representação, foi inserida como propriedade da classe *Scientific Article*.

f) *Illustration* – o conceito do termo ilustração foi utilizado no seu sentido mais amplo, "ação de esclarecer", "breve narrativa ou exemplo que ajuda a compreender ou a esclarecer" (Ilustração, 2017), de forma a ser representada como uma classe que descreve figuras, gráficos, inclusive tabelas, já que esta não se inclui na mesma categoria de figuras e gráficos no contexto dos trabalhos acadêmico-científicos (ABNT, 2012). Essa opção foi adotada com a finalidade de simplificar a representação desses elementos na classe *Scientific Article*.

Foram incluídas ainda as classes referente aos termos:

a) notas (*Note*) – refere-se à comentários externos de utilizadores em relação à classe *Scientific Article*;

b) periódico (*Journal*) – associada à classe *Scientific Article*;

c) editora (*Publisher*) – associada à classe *Journal*;

d) instituição (*Institution*) – está associada à classe *Author* para descrever o vínculo que ele possui com uma instituição ou organização. A Institution também pode ser um autor.

No total, foram definidas 20 classes: *Scientific Article*, *Author*, *Design Research*, *Gene*, *Neoplasm*, *Diagnosis*, *Diagnosis Procedure*, *Imagiology*, *Organism*, *Anatomy*, *Disease*, *Therapeutic*, *Pharmacologic Substance*, *Epidemiology*, *Funding Agency*, *Illustration*, *Note*, *Journal*, *Publisher* e *Institution*.

Os termos identificados como propriedades foram: identificador, volume, número, intervalo de páginas, data de publicação do artigo; título, autoria; resumo; palavras-chave; desenho de investigação, tópico de estudo, ilustrações; referências; conflito de interesses; agência de financiamento associada ao artigo; código do financiamento; apêndices, tipo de artigo; métricas, nível de evidência científica; modo de acesso; direitos autorais; *status*; relação, duração, formato e tamanho do arquivo.

Para a modelagem, foi usado o Unified Modeling Language (UML) por ser de ampla utilização no contexto de modelos de domínios (Bell, 2016, Blacha & Rumbaugh, 2006). Para a elaboração do UML, foi empregado o *software* Microsoft® Office Visio®.

No modelo, foram estabelecidas relações entre as classes, que podem representar um artigo científico. Essas relações descrevem a lógica de relacionamento de todas essas informações, identificando a obrigatoriedade (pelo menos uma associação) e/ou a cardinalidade (a quantidade de instâncias de um objeto que uma classe pode ter em relação à outra). Considera-se que a classe *Scientific Article* pode ter ocorrência de uma ou de várias classes. A exemplo do artigo "*Chronic pancreatitis and pancreatic cancer risk: a systematic review and meta-analysis*" que é representado pela classe *Scientific Articles*, que pode estar associada às classes *Epidemiology*, *Disease*, *Patient* e *Neoplasm* (http://www.nature.com/ajg/journal/vaop/ncurrent/full/ajg2017218a.html?foxtrotcallback =true).

Após a definição das classes, foi especificada a cardinalidade e o mapeamento das propriedades. Os termos mapeados como propriedades em relação às classes priorizaram dois aspectos, o domínio Oncologia e o contexto, que se refere a artigos científicos armazenados em repositórios digitais científicos. O resultado dessas ações é apresentado no diagrama Modelo de Dados (Figura 49).

Vale esclarecer que a cardinalidade entre as classes que representam possíveis tópicos de estudo foi estabelecida da seguinte forma: as classes *Gene*, *Diagnosis*, *Diagnosis Procedure*, *Imagiology*, *Organism*, *Anatomy*, *Disease*, *Therapeutic*, *Pharmacalogic Substance* e *Epidemiology* devem estar associadas a uma e, no máximo, N ocorrências do objeto definido pela classe *Scientific Article*, e a classe *Scientific Article* pode não ter nenhuma ou possuir N ocorrências do objeto definido pelas Classes *Gene*, *Diagnosis*, *Diagnosis Procedure*, *Imagiology*, *Organism*, *Anatomy*, *Disease*, *Therapeutic*, *Pharmacalogic Substance* e *Epidemiology.* Exemplo: as classes *Diagnostic Procedure*, *Pharmacalogic Substance*, *Gene* estão associadas à classe *Scientific Article* (https://www.jstage.jst.go.jp/article/cpb/65/7/65_c17-00098/_pdf).

Em referência à classe *Scientific Article*, deve ter pelo menos uma e no máximo N ocorrências da classe *Neoplasm*. Um exemplo de N ocorrências da classe *Neoplasm* é demonstrado no artigo "*Identication of shared and unique susceptibility pathways among cancers of the lung*, *breast*, *and prostate from genome-wide association studies and tissue-specific protein interactions*" (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4664175). Essa cardinalidade é adotada porque o modelo representa um artigo científico no domínio da Oncologia, não havendo, necessariamente, a obrigatoriedade da associação entre as classes *Gene*, *Diagnosis*, *Diagnosis Procedure*, *Imagiology*,

Organism, *Anatomy*, *Disease*, *Therapeutic*, *Pharmacalogic Substance* e *Epidemiology* e a classe *Scientific Article* (ex.: (http://www.inca.gov.br/ rbc/n_51/v03/pdf/ historia_inca.pdf). A classe *Neoplasm* deve estar associada a pelo menos uma e no máximo N ocorrências da classe *Scientific Article*.

Nesse modelo, destacam-se três classes que possuem características que merecem clarificação:

a) A classe *Disease*, que existe na medida em que outras doenças coexistem com o câncer, ao mesmo tempo em que possui outras propriedades que não as mesmas da classe *Neoplasm*.

b) A classe *Organism* refere-se a uma superclasse maior no domínio da Biologia, na qual se incluem as classes das bactérias, vírus, animais, entre outras. A classe animal possui outras subclasses, entre elas, a do ser humano etc. Pela complexidade do assunto e por esses aspectos ou informações não terem sido abordados por artigos ou pelos estudos apresentados nesta tese, optou-se por incluir a superclasse *Organism*, com atributos da subclasse ser humano, de forma que futuramente sejam trabalhadas outras subclasses, assim como suas propriedades.

A classe *Imagiology* possui atributos próprios, representa as técnicas médicas de diagnóstico por imagem, enquanto que a classe *Illustration* está associada às tabelas, aos gráficos e às figuras. A classe *Imagiology* representa um tópico do artigo, diferentemente da classe *Illustration*, que proporciona a descrição de ilustrações presentes nos artigos, e não a representação de um estudo sobre imagiologia (Ex.: As classes *Neoplasm* e *Imagiology* associadas à classe *Scientific Article*: "*3D reconstruction of cystoscopy videos for comprehensive bladder records*" - https://www.ncbi.nlm.nih.gov/pmc/ articles/PMC5516821/). Assim sendo, a classe *Scientific Article* pode não ter ou ter N ocorrências da classe *Imagiology*, enquanto a classe *Imagiology*, quando existir, deve ter pelo menos uma ocorrência na classe *Scientific Article*.

Conforme já mencionado neste estudo, o mapeamento dos termos (TNIs) para classes e propriedades, foi realizado com base nos seus conceitos e na relação semântica existente com as classes identificadas na taxonomia dos TNIs (Figura 42).

A partir do que foi exposto, foi elaborado o modelo de dados (Figura 50), e na sequência foi desenvolvido o o *Description Set Profile* (DSP) a ser apresentado na próxima subsecção.

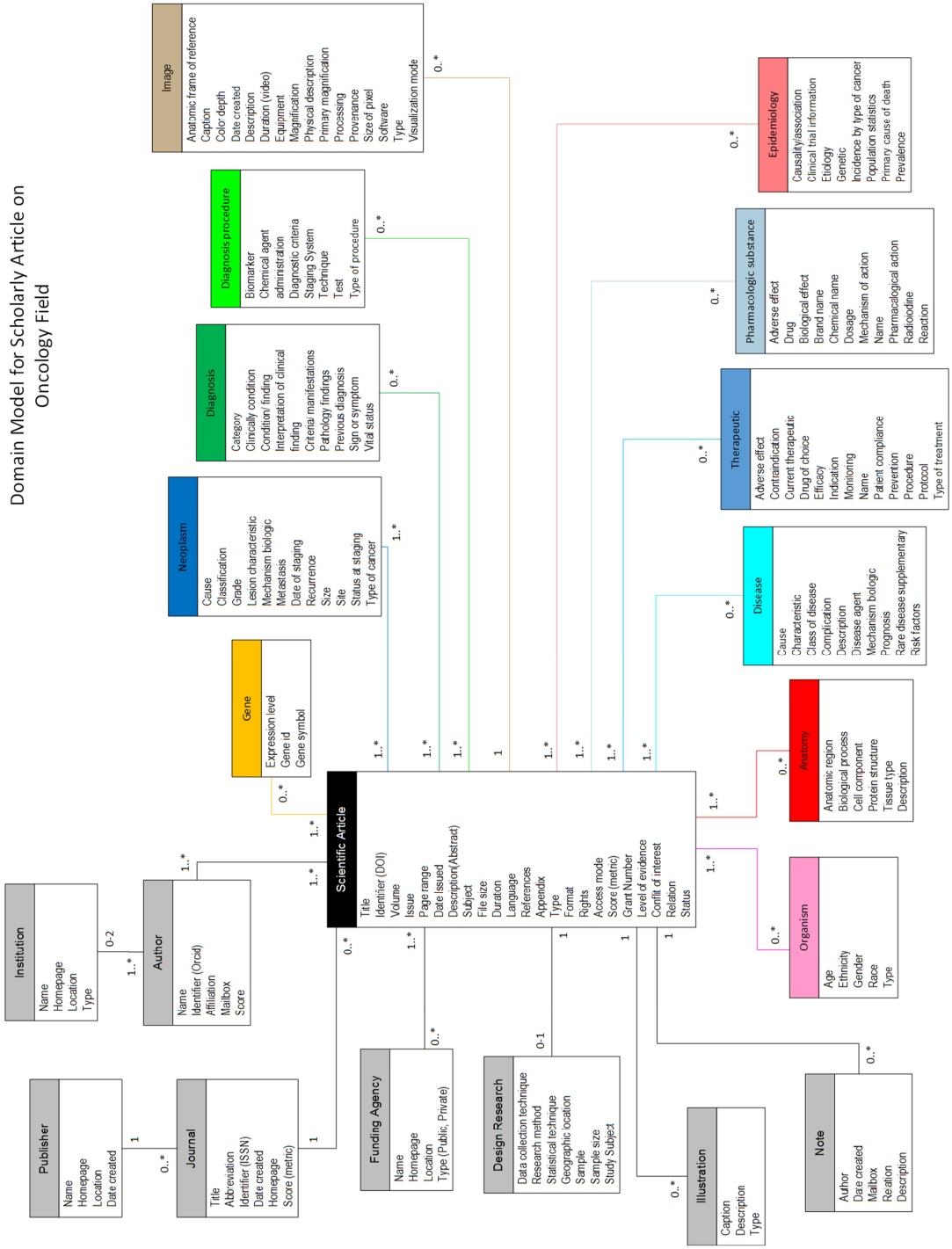


Figura 50. Proposta de modelo de dados para artigos científicos no domínio da Oncologia armazenados em repositórios digitais científicos.

7.3 PRÉ-DESCRIPTION SET PROFILE

O Pré-*Description Set Profile* (DSP)[37] envolveu:

a) desenvolvimento do Modelo de Dados Detalhado – contempla o conjunto de dados validados pelo estudo Delphi;

b) alinhamento de vocabulários – corresponde à tradução dos termos em propriedades de esquemas de metadados. Para tanto, são utilizados os registros de metadados. Neste trabalho, foi adotado, principalmente, o Linked Open Vocabulary (LOV) e foram consultados o Schema.org (health-lifesci.Schema.org) e o Open Metadata Registry;

c) desenvolvimento de um *namespace*, denominado de "onco", para os termos em que não foram identificadas propriedades correspondentes em esquemas ou, quando encontradas, pertenciam a um esquema com pouca ou nenhuma utilização. O *namespace* foi desenvolvido utilizando o esquema RDF (Baptista, 2002) e pode ser visualizado no Apêndice F; e

d) definição da Matriz de Restrições – adotou-se o modelo de Matriz de Restrições sugerido pelo Me4MAP, em que são descritas classes, propriedades e esquemas de codificação. São ainda incluídos os *namespaces*, tipos de restrições e os valores dos *strings* (*Date*, IRI, *String*).

Com base nesse critério, foram selecionados os seguintes esquemas de metadados: dcterms, schema, foaf e rdfs; e as ontologias bibo e dbpedia-owl (Figura 51).

Em relação aos esquemas de metadados Dublin Core (DC e DC Terms), este estudo adota, para a codificação das propriedades do DCMI, as orientações apresentadas pelo DCMI Abstract Model (Powell et al., 2007) e por Rühle, Baker & Johnston (2015).

[37] Expressão utilizada por Malta (2014), ao propor o método Me4MAP para desenvolvimento de perfis de aplicação.

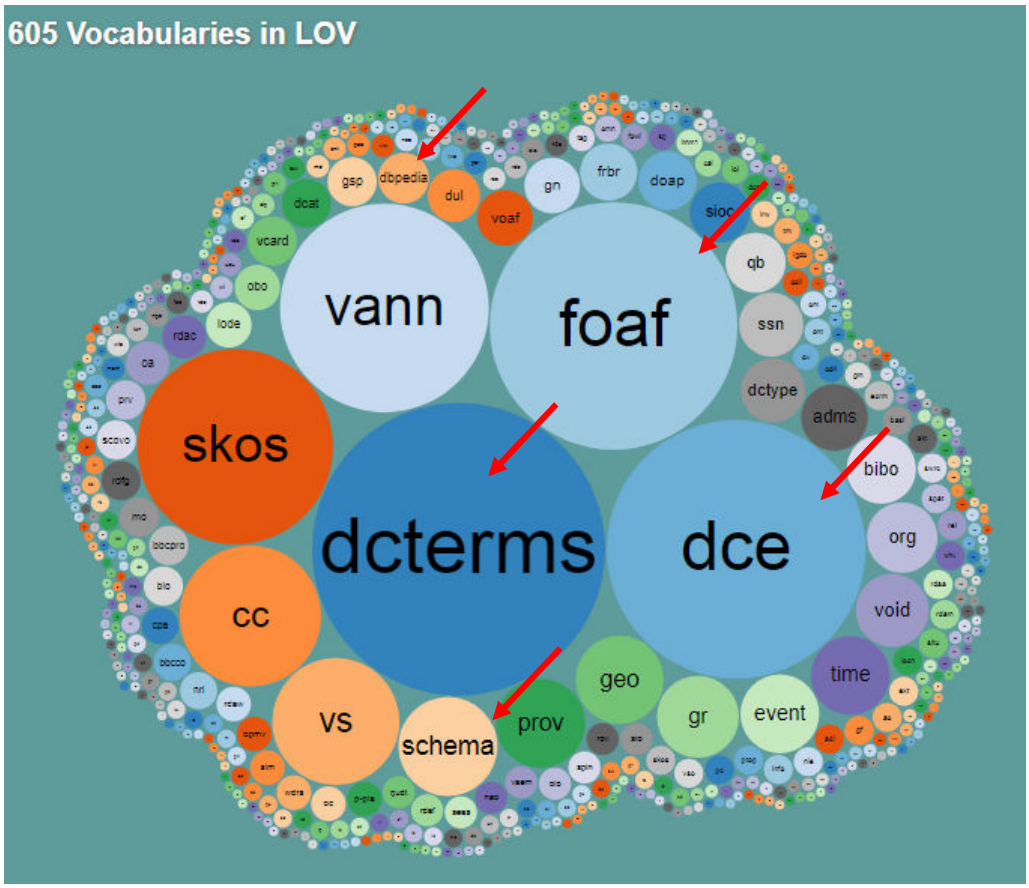


Figura 51. Interface do Language Open Vocabulary (LOV).
Nota: As setas apontam para os vocabulários adotados neste estudo.

A matriz de restrições é apresentada em duas tabelas. A primeira define os *namespaces* utilizados (Tabela 74) e a segunda apresenta as classes e as propriedades (Tabela 75) de acordo com a seguinte organização:

Na primeira linha de cada tabela de descrição consta:

* Description: O nome da Class.
* Classe: O esquema de metadados ou vocabulário, e respectivo termo utilizado (e.g. org:Organization).
* Usage: Uma Description breve do que a Class representa na vida real.

A segunda linha é referente a cada propriedade da tabela de descrição a ser definida:

* Label: A etiqueta dada à propriedade.
* Property: O Esquema de Metadados ou Vocabulário, e respectivo termo utilizado (e.g. dcterms: title).
* Range: O contra-domínio da propriedade: definir se o seu valor é um 'literal' ou um 'non-literal'.
* Value String: No caso do contra-domínio ser um Value String, preencher com 'Sim', caso contrário, preencher com 'Não'.
* SES URI: No caso da propriedade ter um formato específico, indicar o URI desse SES.

- Value URI: No caso do contra-domínio estar restrito a um Value URI, preencher com 'Sim', caso contrário, preencher com 'Não'.
- VES URI: No caso do contra-domínio estar restrito a um VES, indicar o URI.
- Related Description: No caso da possibilidade de a propriedade ser descrita por outras propriedades (e.g. um autor poderá ter como definição o *e-mail*, nome, etc.), deve preencher a lacuna com 'Sim'.

Em relação à cardinalidade, deve conter:

- o valor mínimo O mínimo número de vezes que a propriedade pode ser repetida.
- o valor máximo O máximo número de vezes que a propriedade pode ser repetida.
- Usage: Uma Description breve do que a propriedade representa na vida real (Malta, 2014, p. 114).

Para a definição da matriz de restrição, ainda foram utilizadas as seguintes fontes: Enoksson, F. A MoinMoin Wiki Syntax for Description Set Profiles (s.d.). Retrieved http://dublincore.org/documents/dsp-wiki-syntax/ e Coyle & Baker (2009). Guidelines for Dublin Core application profiles. Retrieved from http://dublincore.org/documents/2009/05/18/profile-guidelines/.

Tabela 74. **Definição dos *namespaces* utilizados.**

Título	**Namespace**	**Prefixo**
Bibliographic Ontology Specification	http://purl.org/ontology/bibo/	bibo
The DBPEDIA Ontology	http://dbpedia.org/ontology/	dbpedia-owl
DCMI Metadata Terms	http://purl.org/dc/terms/	dcterms
Friend of a Friend Vocabulary	http://xmlns.com/foaf/0.1/	foaf
Oncology Vocabulary	http://opendata.dsi.uminho.pt/health/	onco
The RDF Schema Vocabulary	http://www.w3.org/2000/01/rdf-schema#	rdfs
Schema.org vocabular	http://schema.org/	schema.org

Tabela 75. **Matriz de restrições**

Description: Publisher				Class: dbpedia-owl:Publisher			Use: Publishing company			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	rdfs:label	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	No	1	URI	Name of entity responsible for making the resource available.
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0 - 1	URI	A homepage for publisher.
Location	dcterms:coverage	Non literal	Yes	------	Yes	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	No	1	URI	The spatial topic of the resource.
Date created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	------	No	0 - 1	Date	Date of formal issuance of the resource.

Description: Journal				Class: bibo:Journal			Use: A periodical of scholarly journal			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Title	dcterms:title	Literal	Yes	------	No	------	No	1	String	A name given to the journal. Ex.: Advances in cancer research.
Abbreviation	dbpedia-owl:abbreviation	Literal	Yes	NLM Catalog (https://www.ncbi.nlm.nih. gov/pubmed/)	No	------	No	1	String	The abbreviated form representing a resource. Ex.: Adv. cancer res.
Identifier	dcterms:identifier	Literal	Yes	https://www.iso.org/standard/39601.html	No	------	No	1	String	An unambiguous reference to the resource. Ex.: ISSN: 0065-230X
Publisher	dcterms:publisher	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	Yes	1..*	URI	An entity responsible for making the resource available. Ex.: Elsevier, http://id.loc.gov/authorities/names/no95025589
Date Created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	------	No	0-1	Date	Date of formal issuance of the resource. Ex.: 1953-03-23.
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0-1	URI	A homepage for resource. Ex.: http://www.sciencedirect.com/science/bookseries/0065230X.
Score (metric)	onco:score	Non literal	No	------	Yes	------	No	0..*	URI	Represents the metric for resource. Ex.: 82 H Index (http://www.scimagojr.com/journalsearch. php?q =28690&tip=sid)

Description: Scientific Article				Class: schema:ScholarlyArticle			Use: A scholarly academic article, typically published in a journal			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Title	dcterms:title	Literal	Yes	------	No	------	No	1..*	String	A name given to the resource.
Identifier	dc:identifier	Non literal	No	------	Yes	------	No	1..*	URI	An unambiguous reference to the resource within a given context, ex. DOI , Handle
Volume	bibo:volume	Literal	Yes	------	No	------	No	1	String	Volume of a journal.
Issue	bibo:issue	Literal	Yes	------	No	------	No	0-1	String	Issue of a journal.
Page range	bibo:pages	Literal	Yes	------	No	------	No	1	String	Page range of a document.
Date Issued	dcterms:issued	Literal	Yes	http://purl.org/dc/terms/W3CDTF	Yes	------	No	1	Date	Date of publication of the resource.
Author	dcterms:creator.orcid	Non literal	No	------	Yes	https://orcid.org/	Yes	1..*	URI	An entity primarily responsible for making the resource. (Ex.: orcid.org/0000-0002-4055-5882)
Abstract	dcterms:abstract	Literal	Yes	------	No	------	No	1..*	String	A summary of the resource.
Subject	dcterms:subject	Non literal	Yes	------	Yes	http://purl.org/dc/terms/MESH	No	1..*	URI	The topic of the resource.
Language	dcterms:language	Literal	Yes	http://lcweb.loc.gov/standards/iso639-2/langhome.html	No	------	No	1-3	String	A language of the resource. Ex.: eng, por, etc.
References	dc:references	Literal	Yes	------	No	------	No	1	String	A related resource that is referenced, cited, or otherwise pointed to by the described resource.
Appendix	onco:appendix	Non literal	No	------	Yes	------	No	0..*	URI	Text or document prepared by the author
Illustration	onco:illustration	Literal	Yes	------	No	------	Yes	0...*	String	A image, graphic or table within a journal.
Type	dcterms:type	Non literal	No	------	Yes	http://vocabularies.coar-repositories.org/documentation/resource_types/	No	1	URI	The nature or genre of the resource. Ex.: Research, case report, etc.
Format	dcterms:format	Non literal	No	------	Yes	http://www.iana.org/assignments/media-types/	No	1..*	URI	The file format, physical medium, or dimensions of the resource. Ex.: audio, text or video.
Rights	dcterms:rights	Non literal	No	------	Yes	http://creativecommons.org/licenses/by/4.0/	No	1	URI	Information about rights held in and over the resource. Ex.: Licence Creative Commons (CC BY, CC BY-ND, etc.).
Access Mode	onco:accessMode	Non literal	No	------	Yes	http://vocabularies.coar-repositories.org(ongoing implementation)	No	1	URI	Declares the degree of openness of a resource. Ex.: open, embargo or restrict.
Score (metric)	onco:score	Non literal	No	------	Yes	------	No	0..*	URI	Represents the metric being observed. Ex.: Scopus, Scholar Google, Web of Science.
Funding Agency	onco:fundingAgency	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	Yes	1..*	URI	Institutions that offer different forms of financing aimed at conducting research. Ex.: Fundação para a Ciência e a Tecnologia (FTC - Portugal)
GrantNumber	onco:grantNumber	Literal	Yes	——	No	------	No	1..*	String	A data property specifying the grant number of a grant provided by a funding agency. Ex.: UID/Multi/04378/2013
Level of Evidence	schema:evidenceLevel	Non literal	Yes	------	No	http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/	No	1	String	Strength of evidence of the data used to formulate the guideline. Ex.: Level 1A for treatment: a systematic review of randomized clinical trial.
Conflit of Interest	onco:conflitOfInterest	Literal	Yes	------	No	------	No	0-1	String	A situation in which an individual might benefit personally from oficial or professional actions.
Relation	dcterms:relation	Non literal	No	------	Yes	------	No	0..*	URI	A related resource. Ex.: Thesis, Research project.
Status	bibo:status	Non literal	Yes	------	No	http://www.niso.org/publications/rp/RP-8-2008.pdf	No	1	String	Status of a resource. Ex.: Submitted manuscript under review, Accepted manuscript
File size	dcterms:extend	Literal	Yes	------	No	------	No	0-1	String	File size corresponding to the file of article (ex.: 300MB)
Duration	dcterms:extend	Literal	Yes	------	No	------	No	0-1	String	Duration of the audio or video corresponding to the file of article (ex.: 30 min

Description: Author				Class: dcterms:Agent			Use: A resource that acts or has the power to act			
Label	Property	Range	Valuestring	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	foaf:name	Literal	Yes	------	No	------	No	1	String	An entity primarily responsible for making the resource.
Identifier	dc:identifier	Non literal	Yes	------	Yes	------	No	1	URI	An unambiguous reference to the resource within a given context. Ex.: Orcid
Affiliation	schema:affiliation	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	Yes	0..*	URI	An organization that this person is affiliated.
Mailbox	foaf:mbox	Non literal	No	------	Yes	------	No	0-1	mailto:URI	A personal mailbox.
Score (metric)	onco:score	Non iteral	Yes	------	Yes	------	No	0-1	URI	A numeric values measuring performance or quality of something. Ex.: Google Citation.

Description: Institution				Class: dbo: Organization		Use: An organization, establishment or foundation.				
Label	Property	Range	Valuestring	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	rdfs:label	Non literal	Yes	------	Yes	http://id.loc.gov/authorities/names.html	No	1	URI	Name of Institution. Ex.: Universidade do Minho or http://id.loc.gov/rwo/agents/n85096589
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0-1	URI	A homepage for organization.
Location	dcterms:coverage	Non literal	Yes	------	Yes	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	No	1	URI	Physical location of the item.
Type	dcterms:type	Non literal	Yes	------	Yes	https://meshb.nlm.nih.gov/record/ui?ui=D009938	No	0-1	URI	The nature or genre of the resource. Ex.: Academic institution, Hospital, etc.

Description: Funding Agency				Class: onco:FundingAgency		Use: Insitutions or organizations that support scientific or technologic research				
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	rdfs:label	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	No	1	URI	Name of agency (Ex.: Fundação para a Ciência e a Tecnologia, http://id.loc.gov/authorities/names/no00038939
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0..*	URI	A homepage for organization
Location	dcterms:coverage	Non literal	Yes	------	Yes	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	No	1	URI	Physical location of the institution.

Description: Design Research				Class: onco:DesignResearch			Use: Describes how the study performed to answer the research			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Data collection technique	onco:dataCollection	Non literal	Yes	----	Yes	http://www.ddialliance.org/Specification/DDI-CV/ModeOfCollection_1.0.html	No	0..*	URI	Referes to process of gathering data for testing or research purposes.
Research method	onco:researchMethod	Non literal	Yes	----	Yes	http://id.loc.gov/authorities/subjects.html	No	0..*	URI	A method or procedures used in a research.
Statistical technique	onco:statisticalTechnique	Non literal	Yes	----	No	https://stats.oecd.org/glossary/alpha.asp?Let=A	No	0..*	String	A statistical technique used in a research.
Geographic location	dcterms:coverage	Non literal	No	----	Yes	http://purl.org/dc/terms/TGN	No	0..*	URI	Place where the study was developed.
Sample	onco:sample	Literal	Yes	----	No	No	No	1..*	String	The sample of study.
Sample size	onco:sampleSize	Literal	Yes	----	No	No	No	1..*	String	The sample size.
Study Subject	schema:studySubject	Non literal	Yes	----	Yes	https://uts.nlm.nih.gov//snomedctBrowser.html	No	1..*	URI	A subject investigated by the study. Ex.: Therapy, Drug, Diagnosis.

Description: Illustration				Class: onco:Illustration			Use: A numbered and captioned image, graphic or table within a journal.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Caption	schema:caption	Literal	Yes	----	No	----	No	0..*	String	The caption for this object.
Description	dcterms:description	Literal	Yes	----	No	----	No	0..*	String	An account of the resource.
Type	dc:type	Literal	Yes	----	No	----	No	0..*	String	The nature or genre of the resource. Ex.: graphic, table or figure.

Description: Note				Class: bibo:Note		Use: A note about a resource				
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Author	dc:creator	Literal	Yes	----	No	----	Yes	1	String	Na entity primarily responsible for making the resource. (Ex.: Smith, J.)
Date created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	1	Date	Date of creation of the resource
Mailbox	foaf:mbox	Non literal	No	----	Yes	----	No	0-1	mailto: URI	A personal mailbox
Relation	dcterms:relation	Non literal	No	----	Yes	----	No	1	URI	A related resource.
Description	dcterms:description	Literal	Yes	----	No	----	No	1	String	An account of the resource.

Description: Gene				Class: dbpedia-owl:Gene			Use: The functional and physical unit of heredity passed from parente to offspring.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Expression level	onco:proteinExpression	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/projects/CCDS/CcdsBrowse.cgi	No	0-1	URI	The determination of the level of gene transcription in a biological specimen.
Gene id	onco:geneIdentifier	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/gene	No	1	URI	A unique gene name within a specific repository, database, or collection.
Gene symbol	onco:geneSymbol	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/gene	No	1	URI	A unique gene name approved by an organism specific nomenclature committee.

Description: Organism				Class: onco:Organism			Use: A living entity. Ex.: Animal, human being			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Age	dbpedia-owl:age	Literal	Yes	----	No	----	No	0-1	String	Elapsed time since birth.
Ethnicity	dbpedia-owl:ethnicity	Non literal	Yes	----	No	https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html	No	0-1	String	Members share a common history and a sense of identification with the group
Gender	foaf:gender	Non literal	No	----	Yes	http://hl7.org/fhir/v3/AdministrativeGender	No	0-1	URI	The gender of this agent. Ex.: male or female.
Race	dbpedia-owl:race	Non literal	Yes	----	No	https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html	No	0-1	String	An arbitrary classification of a taxonomic group that is a division of a species.
Description	dcterms:description	Non literal	No	----	Yes	https://meshb.nlm.nih.gov/search	No	1..*	URI	Description of a living entity. Ex.: Animal or human being.

Description: Anatomy				Class: onco:Anatomy			Use: Physical structure of animals and plants			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Anatomic region	onco:anatomicRegion	Non literal	No	----	Yes	https://uts.nlm.nih.gov//snomedctBrowser.html	No	1..*	URI	Identifies areas of the body.
Biological process	onco:biologicalProcess	Lliteral	Yes	----	No	----	No	0..*	String	An activity occurring within an organism, between organisms or among organisms and the mechanisms underlying such events.
Cell component	onco:cellComponent	Non literal	No	----	Yes	https://uts.nlm.nih.gov//snomedctBrowser.html	No	0..*	URI	The components of the cell.
Protein structure	onco:proteinStructure	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/structure/	No	0..*	URI	The amino acids and their 3-D arrangement in constituting a protein.
Tissue type	onco:tissueType	Non literal	No	----	Yes	https://uts.nlm.nih.gov//snomedctBrowser.html	No	0..*	URI	Describes collections of differentiated cells, such as epithelium; connective tissue; muscles; nerve tissue.
Description	dcterms:description	Literal	Yes	----	No	----	No	1	String	Describes the referenced anatomy.

Description: Neoplasm				Class: onco:Neoplasm			Use: A benigno or malignant tissue growth resulting from uncontrolled cell proliferation			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Cause	schema:cause	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence.
Classification	onco:classification	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	1	URI	The classification of cancers by histological type; primary site; location where it was first developed.
Date of staging	onco:dateOfStage	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	0-1	Date	Date of he process of defining at what point in the natural history of a malignant disease a patient is when a diagnosis is made.
Grade	onco:gradeTumor	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0-1	URI	A position on a scale of intensity or amount or quality.
Lesion characteristic	onco:lesionCharacteristic	Literal	Yes	----	No	----	No	0..*	String	Characteristic of localized pathological or traumatic structural change, damage, deformity, or discontinuity of tissue, organ, or body part.
Biological mechanism	schema:mechanismBiologic	Non literal	Yes	----	Yes	http://apps.who.int/classifications/icd10/browse/2016/en	No	0..*	URI	Mechanism biologic associated with disease.
Metastasis	onco:metastasis	Non literal	Yes	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0..*	URI	The spread or migration of cancer cells from one part of the body to another.
Recurrence	onco:recurrence	Literal	Yes	----	No	----	No	0-1	String	If the cancer returned.
Site	onco:site	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	2	URI	A position in which cancer develops (histologic and morphologic).
Size	onco:size	Non literal	Yes	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0-1	URI	The physical magnitude of tumor.
Status at staging	onco:statusAtStaging	Non literal	Yes	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0-1	URI	Level of extent of the cancer within the body.
Type of cancer	onco:typeOfCancer	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	1	URI	Type of cancer.

Description: Disease				Class: dbpedia-owl:Disease			Use: A health problem with certain characteristics or symptoms			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Cause	schema:cause	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence.
Characteristic	onco:characteristic	Literal	Yes	----	No	----	No	1..*	String	The distinguishing of a disease aspect.
Class of disease	onco:classOfDisease	Non literal	No	----	Yes	http://www.icd10data.com/	No	1..*	URI	Classification of diseases. Ex.: Breast cancer.
Complication	onco:complication	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	Any disease or disorder that occurs during the course of, or because of, another disease, treatment, or procedure.
Description	dc:description	Literal	Yes	----	No	----	No	1	String	Description of the disease.
Disease agente	onco:diseaseAgent	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	Any vírus, micro-organism, or other substance causing disease.
Mechanism biologic	schema:mechanismBiologic	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	Mechanism biologic associated with disease.
Prognosis	onco:prognosis	Literal	Yes	----	No	----	No	0..*	String	Prognosis.
Rare disease supplementary	onco:rareDisease	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	Rare disease identification.
Risk factors	schema:riskFactor	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	A modifiable or non-modifiable fator that increases the risk of a person contracting a condition. Ex.: Tobacco.

Description: Diagnosis				Class: onco:Diagnosis			Use: The investigation, analysis and recognition of the presence and nature of disease, condition, or injury from expressed signs and symptoms.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Category	schema:category	Non literal	No	----	Yes	http://www.icd10data.com/	No	1	URI	A category for the item. Ex.: Category based on image clinical report – BI-RADS.
Clinically condition	schema:healthCondition	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	Specifying the health condition(s) of a patient, medical study, or other target audience.
Condition/ finding	onco:diseaseEvidence	Non literal	No	----	Yes	https://loinc.org/	No	0..*	URI	Clinical, laboratory or molecular evidence, or lack of evidence of the disease.
Interpretation of clinical finding	onco:interpretingClinicalFinding	Non literal	No	----	Yes	https://loinc.org/	No	0..*	URI	Interpretation of clinical finding.
Criteria/ manifestations	onco:criteriaManifestation	Literal	Yes	----	No	----	No	0-1	String	Criteria based on patient symptoms and signs, test results. Ex.: Amsterdam criteria.
Pathology findings	onco:pathologyFindings	Non literal	No	----	Yes	https://loinc.org/	No	0..*	URI	Pathology findings.
Previous diagnosis	onco: previous Diagnosis	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	Previous diagnosis.
Sign or symptom	schema:signOrSymptom	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	A sign o r symptom of this condition.
Vital status	onco:vitalStatus	Literal	Yes	----	No	----	No	0..*	String	The state or condition of being living or deceased; also includes the case where the vital status is unknown.

Description: Diagnosis procedure				Class:schema: DiagnosticProcedure			Use: A medical procedure intented primarily for diagnostic, as opposed to therapeutic, purposes.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Biomarker	onco:biomaker	Non literal	Yes	----	Yes	https://loinc.org/	No	0..*	URI	Biomarker used.
Chemical agent administration	onco:chemicalAgentAdministration	Non literal	Yes	----	Yes	https://uts.nlm.nih.gov//snomedctBrowser.html	No	0..*	URI	Administration of chemical agent in exams.
Diagnostic criteria	onco:diagnosticCriteria	Literal	Yes	----	No	No	No	0..*	String	A standard from which a judgment concerning a medical diagnosis can be established.
Staging System	onco:stagingSystem	Non literal	No		Yes	http://www.icd10data.com/	No	0..*	URI	A system adoped to describe the extent of cancer in a patient's body. Ex.: TNM Staging System.
Technique	onco:technique	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	Yes	1-..*	URI	Encompass all investigations and tests intended to identify the cause of disease. Ex.: laboratory tests and imaging techniques, Differential diagnosis.
Test	onco:test	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	0..*	URI	Process of determining the disease state of a person. Ex.: cancer mutation test, blood chemistry test, etc.
Type of procedure	schema:procedureType	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1-..*	URI	The type of procedure, for example surgical, noninvasive or percutaneous. Ex.: Clinical diagnosis; computer-aided diagnosis.

Description: Imagiology				Class: schema:ImageTest			Use: Any medical imaging modality typically used for diagnostic purposes			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Anatomic frame of reference	onco:anatomicFrameOfReference	Literal	Yes	----	No	----	No	0..*	String	Image acquisition module.
Caption	schema:caption	Literal	Yes	----	No	----	No	1	String	The caption for this object.
Color depth	onco:colorDepth	Literal	Yes	----	No	----	No	0-1	String	Positive integer indicating the desired minimum total number of bits per color channel used to present a pixel.
Date created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	1	Date	Date of creation of the resource.
Description	dcterms:description	Literal	Yes	----	No	----	No	0..*	String	An account of the image.
Equipment	dbpedia-owl:equipment	Literal	Yes	----	No	----	No	0..*	String	Equipment used to analyze, enhance, interpret or display digital images
Format	dc:format	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	0-1	String	The duration of the item. Ex.: For a video, T19:20+01:00.
Magnification	onco:magnification	Literal	Yes	----	No	----	No	0-1	String	Magnification is related to scaling up visuals or images to be able to see more detail.
Physical description	onco:physicalDescription	Literal	Yes	----	No	----	No	0-1	String	Physical description of the image.
Primary magnification	onco:primaryMagnification	Literal	Yes	----	No	----	No	0-1	String	Initial magnification.
Processing	onco:processingImage	Literal	Yes	----	No	----	No	0...*	String	The use of data processing systems to analyze, enhance, interpret or display digital images.
Provenance	dcterms:provenance	Non literal	Yes	----	No	----	No	1	URI	Ownership and custody of the resource of since its creation. Ex.: Hospital, Laboratory.
Size of pixel	onco:sizeOfPixel	Literal	Yes	----	No	----	No	0-1	String	Image size expressed in pixels.
Software	onco:software	Literal	Yes	----	No	----	No	0-1	String	A computer program in source of compiled form.
Type	dcterms:type	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1*	URI	The nature or genre of the resource. Ex.: radiography, videolaryngoscopy.
Visualization Mode	onco:visualizationMode	Literal	Yes		No	----	No	0-1	String	Options for visualization (ex.: 3D image).

Description: Therapeutic				Class: onco:Therapeutic			Use: Relating to or involved in therapy			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Adverse effect	onco:adverseEffect	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	An unexpected medical problem that happens during treatment with a drug or other therapy.
Contraindication	schema:contraindication	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	A contraindication for this therapy
Current therapeutic	onco:currentTherapeutic	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	0..*	URI	The name of the therapy that is currently being administered.
Drug of choice	schema:drugClass	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	The class of drug this belongs to (ex., statins)
Efficacy	onco:efficacy	Literal	Yes	----	No	----	No	0..*	String	The ability of a chemical agent or a device to treat a disease.
Indication	schema:indication	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	0..*	URI	A fator that indicates use of this therapy for treatment and/or prevention of a condition, symptom, etc.
Monitoring	onco:monitoring	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	0..*	URI	How a person or a condition is assited or verified.
Patient compliance	onco:patientCompliance	Literal	Yes	----	No	----	No	0..*	String	The act of following a medical regimen or schedule correctly and consistently, including taking medicines or following a diet.
Prevention	onco:prevention	Literal	Yes	----	No	----	No	0..*	String	Describes methods of preventing cancer including limit of exposure to carcinogens and infectious agents, diet, education, etc
Procedure	onco:medicalProcedure	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1..*	URI	Describes a treatment or action taken to prevent or treat disease
Protocol	onco:protocol	Literal	Yes	----	No	----	No	0..*	String	Describes a rule which guides how an activity should be performed.
Type of treatment	onco:typeOfTreatment	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1..*	URI	Cancer treatment adopted to non directed or directed cancer Ex.: control, remove or destroy cancer or prolong a patient's live, alleviate pain.

Description: Pharmacologic Substance				Class: onco:PharmacologicSubstance			Use: Any substance, other than food, that is used to prevent, diagnose, treat or relieve symptoms of a disease or abnormal condition			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Adverse effects	onco:adverseEffect	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	An unexpected medical problem that happens during treatment with a drug or other therapy.
Drug	onco:drug	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	Specifying a drug or medicine used in a medication procedure
Biological effects	onco:biologicalEffect	Non literal	Yes	----	No	http://www.icd10data.com/	No	0..*	URI	Biological effect that the any product provide.Ex.: Radiation
Brand name	dbpedia-owl:brand	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	Brand name. Ex.: Ibuprofen 400 to Ibuprofen.
Chemical name	onco:chemicalName	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	Chemical name. Ex.: Ibuprofen; Brufen to Ibuprofen
Dosage	schema:doseValue	Literal	Yes	----	No	----	No	0..*	String	The value of the dose. Ex. 500 mg
Mechanism of action	schema:mechanismOfAction	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	0..*	URI	The specific biochemical interaction through which this drug or supplement produces its pharmacological effect
Pharmacalogical action	onco:pharmacologicAction	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	0..*	URI	Actions of drugs and their effects on living organisms
Radioiodine	onco:radioiodine	Literal	Yes	----	No	----	No	0..*	String	Use of radioiodine.
Reaction	onco:reaction	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	A description of the action in response to some kind of stimulus.

Description: Epidemiology				Class: onco:Epidemiology			Use: The study of the patterns, causes, and control of disease in groups of people			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Causality/ association	onco:causality	Literal	Yes	----	No	----	No	0..*	String	Relationship exists between a cause or effect. Ex.: radiation.
Clinical trial information	onco:clinicalTrialInformation	Literal	Yes	---=p	No	----	No	0-1	String	Report of a pre-planned medical or veterinary clinical case study involving a randomized, evidence-controlled clinical trial.
Etiology	onco:etiology	Non literal	Yes	----	No	http://www.icd10data.com/	No	0..*	URI	The cause of a disease or abnormal condition.
Genetic	onco:geneticEpidemiology	Literal	Yes	----	No	----	No	0-1	String	Interdisciplinary clinical, epidemiologic, genetic, and laboratory studies of persons, families, and populations at high risk of cancer.
Incidence by type of câncer	onco:incidence	Literal	Yes	----	No	----	No	1..*	String	The relative frequency of occurrence of cancer.
Population statistics	onco:populationStatistics	Literal	Yes	----	No	----	No	0..*	String	A value, which represents a particular population characteristic.
Prevalence	onco:prevalence	Literal	Yes	----	No	----	No	0..*	String	The total number of cases of a given disease in a specified population at a designated time.
Primary cause of death	onco:primaryCauseofDeath	Non literal	Yes	----	No	http://www.icd10data.com/	No	0-1	URI	The first significant event which ultimately led to death.

Faz parte dessa matriz de restrições o esquema de metadados Onco, desenvolvido com a finalidade de contemplar classes e propriedades não identificadas em outros esquemas ou, quando identificadas, apresentavam número reduzido de utilizadores ou, ainda, a utilização deles proporcionava uma diversidade de *namespaces* muito grande (Baptista, 2002).

O esquema de metadados Onco foi desenvolvido e o *namespace* foi registrado no domínio da Universidade do Minho. Para definir as classes e as propriedades do referido esquema, foram utilizados os sistemas de organização do conhecimento NCI Thesaurus e SNOMED, ontologias (OBO, schema.org), dicionários e livros.

A definição dos VES e dos SES foi baseada na:

a) seleção de regras e vocabulários resultantes dos estudos desenvolvidos no âmbito desta tese (Tabela 36);
b) identificação de vocabulários controlados e/ou regras de sintaxe por meio de buscas utilizando o Google;
c) consulta a *sites* de instituições/órgãos que desenvolvem atividades no domínio da Oncologia (ex.: National Cancer Institute; Serviço Nacional de Saúde, Portugal; Instituto Nacional de Oncologia - INCA, Brasil, serviços de saúde dos Estados Unidos, NISO).

Os vocabulários identificados foram analisados considerando o seu objetivo, a representatividade de quem os utiliza e a atualização do vocabulário. Salienta-se que não foi avaliada a qualidade dos vocabulários por não ser objeto deste estudo.

A Tabela 76 apresenta os esquemas de codificação de vocabulário e a Tabela 77, os esquemas de codificação de sintaxe adotados no Onco-MAP, com as respectivas descrições e IRI para acesso.

Tabela 76. **Esquemas de codificação de vocabulários adotados no Onco-MAP**

Vocabulário	Descrição	IRI
CD-10-PCS	Is a free reference website designed to code and classify procedures associated with hospital used in the United States.	https://www.cms.gov/Medicare/Coding/ICD10/index.html
COAR controlled vocabularies	Is a set of controlled vocabularies for the bibliographic metadata elements used in records describing	http://vocabularies.coar-repositories.org
DDI Controlled Vocabulary for Mode of Collection	The procedure, technique, or mode of inquiry used to attain the data.	http://www.ddialliance.org/Specification/DDI-CV/ModeOfCollection_1.0.html

Vocabulário	Descrição	IRI
Gene (NCBI)	Gene integrates information from a wide range of species. A record may include nomenclature, Reference Sequences (RefSeqs), maps, pathways, variations, phenotypes, and *links* to genome-, phenotype-, and locus-specific resources worldwide	https://www.ncbi.nlm.nih.gov/gene
Getty Thesaurus of Geographic Names® Online	Is a structured vocabulary of geographic names for indexing art and architecture.	http://www.getty.edu/research/tools/vocabularies/tgn/index.html
Glossary of Statistical Terms (OECD)	Contains a comprehensive set of definitions of the main data items collected by the Organisation.	https://stats.oecd.org/glossary/alpha.asp?Let=A
ICD-O-3	Is a dual classification, with coding systems for both topography and morphology.	http://codes.iarc.fr/usingicdo.php
International Classification of Diseases Ninth Revision Clinical Modification (ICD-9-CM)	Use to code and classify morbidity data from inpatient and outpatient records, physician office, andd most National Center for Health Statistics (NCHS) *survey*s. ICD-9-CM is the oficial system of assigning codes to diagnosis and procedures associated with hospital utilization in the United States.	http://icd9cm.chrisendres.com/index.php (or http://www.icd10data.com/)
International Classification of Diseases Tenth Revision (ICD-10)	Used to code and classify mortality data for reporting to the World Health Organization (WHO)	http://apps.who.int/classifications/icd10/browse/2010/en
Journal Article Versions (JAV): recomendations of the NISO/ALPSP JAV Technical Working Group (NISO)	Provide a simple, practical way of describing the versions of scholarly journal articles that typically appear online before, during, and after formal journal publication.	http://www.niso.org/publications/rp/RP-8-2008.pdf
Library of Congress Subject Headings (LCSH)	Provide subject access to their collections	http://id.loc.gov/authorities/subjects.html
Licença Creative Commons	Is used when na author wants to give people the right to share, use, and build upon a work that they have created.	http://creativecommons.org/licenses/by/4.0/
Logical Observation identifiers Names and Codes (LOINC) Terminology	Provides a standard set of universal names and codes for identifying individual laboratory and clinical results.	https://loinc.org/
Media Types	Specifies that Media Types (formerly known as MIME types)	http://www.iana.org/assignments/media-types/

Vocabulário	Descrição	IRI
Medical Subject Heading	Is the NLM controlled vocabulary thesaurus used for indexing articles for PubMed	https://meshb.nlm.nih.gov/search
NCI Thesaurus	An ontology-like vocabulary that includes broad coverage of the cancer domain, including cancer related diseases, findings and abnormalities; anatomy; agents, drugs and chemicals; genes and gene products and so on. In certain areas, like cancer diseases	https://ncit.nci.nih.gov/ncitbrowser/
Nomenclatura Sistematizada de Medicina – Termos Clínicos (SNOMED) –	Designated data standards for use in the electronic exchange of clinical health information. It is considered to be the most comprehensive, multilingual clinical healthcare terminology in the world	https://www.nlm.nih.gov/healthit/snomedct/index.html
ORCID Connecting Research and Researchers	Provides a persistent digital identifier that distinguishes you from every other researcher	https://orcid.org/
Race and Ethnic Standards for Federal Statistics and Administrative Reporting (Centers for Disease Control and Prevention)	Provides standard classifications for record keeping, collection, and presentation of data on race and ethnicity in Federal program administrative reporting and statistical activities	https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html
RxNorm – National Library of Medicine - NIH	RxNorm is a normalized naming system for generic and branded drugs; and a tool for supporting semantic interoperation between drug terminologies and pharmacy knowledge base systems.	https://mor.nlm.nih.gov/RxNav/
Standards for the Classification of Federal Data on Race and Ethnicity	These standards are used governmentwide for recordkeeping, collection, and presentation of data on race and ethnicity in Federal statistical activities and program administrative reporting	https://www.whitehouse.gov/omb/fedreg_notice_15
The CEBM 'Levels of Evidence 1'	Specifies the levels of clinical evidence from a study	http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/
The Consensus CDS (CCDS)	Identify a core set of protein coding regions	https://www.ncbi.nlm.nih.gov/projects/CCDS/CcdsBrowse.cgi
The Library of Congress Name Authority File (NAF	Provides authoritative data for names of persons, organizations, events, places, and titles	http://id.loc.gov/authorities/names.html
v3 Code System AdministrativeGender (Health Level Seven International)	The gender of a person used for adminstrative purposes (as opposed to clinical gender)	http://hl7.org/fhir/v3/AdministrativeGender

Tabela 77. **Esquemas de codificação de sintaxe adotados no Onco-MAP**

Esquema de Sintaxe	Descrição	IRI
ISO 639.2	Codes for the representation of names of languages	http://lcweb.loc.gov/standards/iso639-2/langhome.html
Date and time format (W3C)	Describes a large number of date/time formats	http://purl.org/dc/terms/W3CDTF
ISO 3297:2007 – Information and documentation – International standard serial number (ISSN)	Defines and promotes the use of a standard code (ISSN) for the unique identification of serials and other continuing resources	https://www.iso.org/standard/39601.html
NLM Catalog	The NLM Catalog provides access to NLM bibliographic data for journals, books, audiovisuals, computer software, electronic resources and other materials	NLM Catalog (https://www.ncbi.nlm.nih. gov/pubmed/)

SÍNTESE DA SECÇÃO 7

A proposta *Linked Open Data* é um paradigma disruptivo que pode impactar positivamente a descoberta e a produção de conhecimento que alimenta os avanços científicos no tratamento do câncer. No entanto, requer ainda muito esforço quanto ao uso de padrões e boas práticas referentes aos metadados, sendo o PAM uma das alternativas para consolidação, de acordo com a comunidade de metadados.

Nesta Seção foi apresentado os procedimentos adotados para o desenvolvimento do perfil de aplicação de metadados – Onco-MAP. Trata-se de um perfil voltado para artigos científicos no contexto dos repositórios digitais científicos no domínio da Oncologia. O Onco-MAP foi elaborado seguindo a lógica do Me4MAP, utilizando sistemas de organização do conhecimento como o SNOMED e o NCI Thesaurus para definição dos termos; e os registros de metadados, LOV, Metadata Registry e Schema.org,,para identificação das propriedades.

Para definição dos esquemas de codificação, recorreu à literatura, ao resultado do *Survey* aplicado aos gestores de RDCs, às organizações que regulamentam atividades associadas à Saúde, como ministérios/secretarias, e às buscas no Google.

A próxima Seção apresenta Onco-MAP e os procedimentos adotados para a sua validação.

8 ONCO-MAP: PERFIL DE APLICAÇÃO DE METADADOS

Nesta Seção, é apresentada a concretização do objetivo deste estudo: o desenvolvimento de um Perfil de Aplicação Dublin Core para artigos científicos no contexto dos RDCs no domínio da Oncologia, o qual foi denominado de Onco-MAP, de acordo com princípios *Linked Open Data* (LOD) e baseado nos TNIs de médicos e investigadores que atuam nesse domínio.

O *Description Set Profile* vem ao final da Seção, após a narrativa dos procedimentos adotados para a sua validação, realizada a partir do *Pré-Description Set Profile* exposto na Subseção 7.3

8.1 VALIDAÇÃO DO PERFIL DE APLICAÇÃO DE METADADOS – ONCO-MAP

A proposta da validação foi obter consenso sobre o modelo de dados, nomeadamente em relação às atribuições das classes e das propriedades. As questões apresentadas no questionário buscaram identificar: se as classes representam um artigo científico, se as propriedades estão devidamente representadas nas classes ou se alguma delas estaria mais bem representada em outra classe; se havia casos de ambiguidade (pode ter mais do que um sentido ou significado) ou redundância (termos distintos para propriedades, porém com o mesmo significado e, portanto, o mesmo conteúdo); se as associações entre as classes (que complementam o conjunto de informações) estão adequadamente representadas no que tange à obrigatoriedade (nenhuma ou pelo menos uma associação) e à cardinalidade (a quantidade de instâncias de um objeto que uma classe pode ter em relação à outra).

A validação do modelo, que envolve a ratificação das classes e propriedades, foi realizada por bibliotecárias(os) que exercem atividades no domínio da Oncologia. Optou-se por validar com os profissionais de informação em virtude do papel que eles desempenham como intermediários no processo de comunicação científica, nas fases de recuperação e disseminação da informação e no processo de representação temática da informação.

A escolha por validar o PAM com os(as) bibliotecários(as) se deu pela própria característica deste trabalho, que está voltado para a concepção de um artefacto desenvolvido por profissional da informação; que utiliza informações obtidas de médicos e investigadores da área de Oncologia (TNIs); e que se refletirá na estruturação e na organização dos sistemas de informação que visam atender às necessidades informacionais dos atores que atuam nesse domínio. Dessa forma, o que aqui está posto é a validação do que a autora produziu, o Onco-MAP. O Delphi foi a técnica utilizada para o realização dessa atividade.

A técnica Delphi tem sido utilizada largamente em diferentes domínios: Educação, Psicologia, Medicina Ciência da Informação, Administração e Sistemas de Informação (Batista & Cunha, 2007, Okoli & Pawlowski, 2004, Soares, 2010, Wilson, 1971). Assim como outras técnicas, não existe consenso quanto: ao número de participantes, ao número de rondas e, particularmente, ao índice de concordância, já que essa técnica busca o consenso entre os especialistas.

O estabelecimento do número de participantes e de rondas depende da composição do grupo, do objetivo do estudo e dos critérios estabelecidos pelo investigador (Day & Bobeva, 2005, Fink, Kosecoff, Chassin, & Brook, 1984, Passos, Sequeira, & Fernandes, 2013, Powell, 2013). Esse estudo foi realizado com profissionais bibliotecários associados a instituições de reconhecida relevância na área da Oncologia, totalizando o envio de 28 *e-mails* e a distribuição de 10 *links* (opção necessária em razão do controle de segurança das instituições). Partindo do princípio de que alguns autores preconizam que o número de participantes é definido pelo investigador, e não existindo uma orientação rígida quanto a esse número (Day & Bobeva, 2005, Fink, Kosecoff, Chassin, & Brook, 1984, Passos, Sequeira, & Fernandes, 2013, Powell, 2013), a autora optou por priorizar a qualificação e a competência dos integrantes em relação ao domínio e ao compromisso de participar do estudo (Powell, 2013).

O número de rondas foi definido pelo investigador com base na confiabilidade dos resultados da segunda ronda (Kayo & Securato, 1997, Murry & Hammons, 1995). De modo geral, são utilizadas de duas a três rodadas (Keeney, Hasson & McKenna, 2001, Murry & Hammons, 1995).

A avaliação de concordância, ou medida de consenso, pode ser arbitrada pelo investigador. Na literatura, observam-se os seguintes índices: concordância igual ou superior a 2/3 das respostas (Delbecq, Van de Ven & Gustaffson, 1986); ou ainda critérios defendidos por Fink, Kosecoff, Chassin, & Brook (1984, p. 981):

- No final da votação, qualquer tópico ou assunto aceitos por pelo menos X% dos participantes são adotados;
- Após cinco rodadas de votação, o número X dos tópicos recebeu a maioria dos votos aprovados;
- Todos os tópicos são classificados em uma escala de 1 a 5. Somente aqueles tópicos que recebem uma classificação média de 3.0 ou superior são aceitos;
- Todos os tópicos são classificados em uma escala de 1 em 3. Todos os tópicos recebendo uma classificação de '1' de pelo menos 51% de os participantes são adotados;
- Qualquer tópico é descartado se for fortemente oposto por pelo menos X% dos participantes.

Nesse sentido, este estudo considera que o critério para determinação do consenso considerará que, na última ronda, a propriedade deverá obter 75% de concordância, índice preconizado pelos autores Murry & Hammons (1995). A partir do estabelecimento desses critérios, foi dado início ao estudo Delphi, descrito a seguir.

8.1.1 Estudo Delphi - primeira ronda

A primeira ronda visou a sistematizar os resultados das questões, com o objetivo de buscar consenso entre os participantes. Nessa perspectiva, as questões foram analisadas em relação ao nível de concordância sobre as classes, as propriedades e o modelo.

Amostra

Após as alterações decorrentes do pré-teste, o questionário foi enviado por *e-mail* e por *link* para bibliotecários(as) selecionados(as) com base nos seguintes critérios: experiência profissional superior a cinco anos, desempenho de função em bibliotecas ou centros de investigação na área da Oncologia ou da Medicina (nesse caso, após confirmação da existência de atividades relacionadas com a Oncologia).

A primeira ronda foi realizada no período de 15-08-2017 a 11-09-2017. Contou com a participação de oito profissionais (bibliotecários e bibliotecárias), o que corresponde a 38% de respostas do total de questionários enviados (22 *e-mails* válidos). Nessa amostra, não foram incluídos os respondentes do pré-teste. Do total de 22 *e-mails*, dois foram devolvidos e dois rejeitados. Dos nove participantes que aceitaram o termo de consentimento, um não respondeu às questões por não estar mais atuando na área.

Questionário

A primeira ronda do Estudo Delphi utilizou um questionário estruturado em 27 perguntas, em que 24 são de múltipla escolha com a seguinte estrutura: 1 questão referente ao termo de consentimento; 1 para inclusão de dados pessoais; 1 imagem do modelo; 22 questões de múltipla escolha em que constam classes e propriedades com seus respectivos conceitos e com campo para observações; 1 em que é solicitada a opinião sobre o modelo; e 2 para inclusão de informações para contato (APÊNDICE D).

Foi solicitado aos bibliotecários(as) que indicasse, de acordo com cada classe, se a propriedade: deveria ser mantida como propriedade; se não deveria ser mantida como propriedade; se estava associada à classe; ou se não sabia.

Resultado

Questão P1. Refere-se ao Termo de consentimento que todos os respondentes deveriam aceitar para participar do inquérito. Oito participantes confirmaram a participação.

Questão P2. Perfil dos profissionais

Os profissionais que participaram do estudo Delphi são bibliotecárias(os), exercem atividades de referência, catalogação e/ou gestão em bibliotecas universitárias ou hospitalares, com experiência entre 10 e 30 anos; e possuem especialização ou mestrado.

Da questão P3 à questão P22, foi solicitado ao respondente que se pronunciasse sobre a representatividade das classes e suas propriedades para a descrição de artigos científicos:

a) se as propriedades estavam devidamente representadas nas classes

b) se alguma delas estaria mais bem representada em outra classe; e

c) se havia ambiguidade ou redundância, neste caso a propriedade seria eliminada.

No questionário, todas as classes e propriedades possuíam definições, embora não constem das Tabelas 78 a 96. As definições incluídas no questionário podem ser visualizadas na matriz de restrições ou no Apêndice D. As Tabelas apresentam as propriedades e o percentual de concordância entre os respondentes.

Tabela 78. Questão P3. Classe: **Publisher**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name	100,00%	0,00%	0,00%	0,00%
Homepage	87,50%	12,50%	0,00%	0,00%
Location	87,50%	12,50%	0,00%	0,00%
Type	87,50%	12,50%	0,00%	0,00%

Tabela 79. Questão P4. Classe: **Journal**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Title	100,00%	100,00%	0,00%	0,00%
Abbreviation	87,50%	12,50%	0,00%	0,00%
Identifier	100,00%	0,00%	0,00%	0,00%
Publisher	100,00%	0,00%	0,00%	0,00%
Date Created	100,00%	0,00%	0,00%	0,00%
Homepage	100,00%	0,00%	0,00%	0,00%
Score	100,00%	0,00%	0,00%	0,00%

Tabela 80. Questão P5. Classe: **Scientific Article**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Title	100,00%	0,00%	0,00%	0,00%
Identifier	100,00%	0,00%	0,00%	0,00%
Volume	100,00%	0,00%	0,00%	0,00%
Issue	100,00%	0,00%	0,00%	0,00%
Page range	87,50%	12,50%	0,00%	0,00%
Date Issued	100,00%	0,00%	0,00%	0,00%
Author	100,00%	0,00%	0,00%	0,00%
Abstract	100,00%	0,00%	0,00%	0,00%
Subject	87,50%	12,50%	0,00%	0,00%
Language	87,50%	12,50%	0,00%	0,00%
References	87,50%	12,50%	0,00%	0,00%
Appendix	62,50%	37,50%	0,00%	0,00%
Illustration	75,00%	25,00%	0,00%	0,00%
Type	75,00%	25,00%	0,00%	0,00%
Format	87,50%	0,00%	12,50%	0,00%
File size	62,50%	12,50%	12,50%	12,50%
Duration	50,00%	25,00%	12,50%	12,50%
Rights	7500%	25,00%	0,00%	0,00%
Access Mode	87,50%	12,50%	0,00%	0,00%
Score	87,50%	12,50%	0,00%	0,00%
Funding Agency	62,50%	25,00%	0,00%	12,50%
GrantNumber	62,50%	37,50%	0,00%	0,00%
Level of Evidence	100,00%	0,00%	0,00%	0,00%
Conflit of Interest	75,00%	25,00%	0,00%	0,00%
Relation	75,00%	25,00%	0,00%	0,00%
Status	75,00%	25,00%	0,00%	0,00%

Tabela 81. Questão P6 . Classe: **Author**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name	100,00%	0,00%	0,00%	0,00%
Identifier	87,50%	12,50%	0,00%	0,00%
Affiliation	87,50%	12,50%	0,00%	0,00%
Mailbox	100,00%	0,00%	0,00%	0,00%
Score	87,50%	0,00%	0,00%	12,50%

Tabela 82. Questão P7. Classe: **Institution**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name	100,00%	0,00%	0,00%	0,00%
Homepage	87,50%	12,50%	0,00%	0,00%
Location	87,50%	12,50%	0,00%	0,00%
Type	87,50%	12,50%	0,00%	0,00%

Tabela 83. Questão P8. Classe: **Illustration**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Caption	75,00%	25,00%	0,00%	0,00%
Description	62,50%	37,50%	0,00%	0,00%
Type	75,00%	25,00%	0,00%	0,00%

Tabela 84. Questão P10. Classe: **Note**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Author	62,50%	37,50%	0,00%	0,00%
Date created	62,50%	37,50%	0,00%	0,00%
Mailbox	62,50%	37,50%	0,00%	0,00%
Relation	75,00%	25,00%	0,00%	0,00%
Description	62,50%	37,50%	0,00%	0,00%

Tabela 85. Questão P11. Classe: **Design Research**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Data collection technique	87,50%	12,50%	0,00%	0,00%
Research method	100,00%	0,00%	0,00%	0,00%
Statistical technique	100,00%	0,00%	0,00%	0,00%
Geographic location	87,50%	12,50%	0,00%	0,00%
Sample	75,00%	12,50%	0,00%	12,50%
Sample size	75,00%	12,50%	0,00%	12,50%
Study Subject	100,00%	0,00%	0,00%	0,00%

Tabela 86. Questão P12. Classe: **Gene**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Expression level	87,50%	12,50%	0,00%	0,00%
Gene id	87,50%	12,50%	0,00%	0,00%
Gene symbol	87,50%	12,50%	0,00%	0,00%

Tabela 87. Questão P13. Classe: **Organism**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Age	87,50%	12,50%	12,50%	0,00%
Ethnicity	100,00%	0,00%	0,00%	0,00%
Gender	100,00%	0,00%	0,00%	0,00%
Race	87,50%	12,50%	12,50%	0,00%
Description	100,00%	0,00%	0,00%	0,00%

Tabela 88. Questão P14. Classe: **Anatomy**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Anatomic region	100,00%	0,00%	0,00%	0,00%
Biological process	87,50%	12,50%	0,00%	0,00%
Cell component	87,50%	12,50%	0,00%	0,00%
Protein structure	87,50%	12,50%	0,00%	0,00%
Tissue type	100,00%	0,00%	0,00%	0,00%
Description	87,50%	12,50%	0,00%	0,00%

Tabela 89. Questão P15. Classe: **Neoplasm**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Cause	87,50%	12,50%	0,00%	0,00%
Classification	100,00%	0,00%	0,00%	0,00%
Date of staging	87,50%	12,50%	0,00%	0,00%
Grade	87,50%	12,50%	0,00%	0,00%
Lesion characteristic	100,00%	0,00%	0,00%	0,00%
Biological mechanism	100,00%	0,00%	0,00%	0,00%
Metastasis	100,00%	0,00%	0,00%	0,00%
Recurrence	100,00%	0,00%	0,00%	0,00%
Site	87,50%	12,50%	0,00%	0,00%
Size	87,50%	12,50%	0,00%	0,00%
Status at staging	100,00%	0,00%	0,00%	0,00%
Type of cancer	100,00%	0,00%	0,00%	0,00%

Tabela 90. Questão P16. Classe: **Disease**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Cause	100,00%	0,00%	0,00%	0,00%
Characteristic	100,00%	0,00%	0,00%	0,00%
Class of disease	100,00%	0,00%	0,00%	0,00%
Complication	100,00%	0,00%	0,00%	0,00%
Description	87,50%	0,00%	0,00%	12,50%
Disease agent	100,00%	0,00%	0,00%	0,00%
Biological mechanism	100,00%	0,00%	0,00%	0,00%
Prognosis	100,00%	0,00%	0,00%	0,00%
Rare disease supplementary	87,50%	12,50%	0,00%	0,00%
Risk factors	100,00%	0,00%	0,00%	0,00%

Tabela 91. Questão P17. Classe: **Diagnosis**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Category	87,50%	12,50%	0,00%	0,00%
Clinically condition	87,50%	12,50%	0,00%	0,00%
Condition/ finding	87,50%	12,50%	0,00%	0,00%
Interpretation of clinical finding	87,50%	12,50%	0,00%	0,00%
Criteria/ manifestations	87,50%	12,50%	0,00%	0,00%
Pathology findings	100,00%	0,00%	0,00%	0,00%
Previous diagnosis	87,50%	12,50%	0,00%	0,00%
Sign or symptom	75,00%	25,00%	0,00%	0,00%
Vital status	87,50%	12,50%	0,00%	0,00%

Tabela 92. Questão P18. Classe: **Diagnostic procedure**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Biomarker	87,50%	12,50%	0,00%	0,00%
Chemical agent administration	100,00%	0,00%	0,00%	0,00%
Diagnostic criteria	87,50%	12,50%	0,00%	0,00%
Staging system	87,50%	12,50%	0,00%	0,00%
Technique	87,50%	12,50%	0,00%	0,00%
Test	100,00%	0,00%	0,00%	0,00%
Type of procedure	100,00%	0,00%	0,00%	0,00%

Tabela 93. Questão P19. Classe: **Imagiology**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Anatomic frame of reference	87,50%	12,50%	0,00%	0,00%
Caption	87,50%	12,50%	0,00%	0,00%
Color depth	62,50%	37,50%	0,00%	0,00%
Date created	87,50%	12,50%	0,00%	0,00%
Description	87,50%	12,50%	0,00%	0,00%
Equipment	75,00%	25,00%	0,00%	0,00%
Format	75,00%	25,00%	0,00%	0,00%
Magnification	87,50%	12,50%	0,00%	0,00%
Physical description	75,00%	25,00%	0,00%	0,00%
Primary magnification	50,00%	25,00%	0,00%	25,00%
Processing	75,00%	25,00%	0,00%	0,00%
Provenance	87,50%	12,50%	0,00%	0,00%
Size of pixel	62,50%	37,50%	0,00%	0,00%
Software	75,00%	25,00%	0,00%	0,00%
Type	87,50%	12,50%	0,00%	0,00%
Visualization Mode	87,50%	12,50%	0,00%	0,00%

Tabela 94. Questão P20. Classe: **Epidemiology**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Causality/ association	100,00%	0,00%	0,00%	0,00%
Clinical trial information	87,50%	0,00%	12,50%	0,00%
Etiology	100,00%	0,00%	0,00%	0,00%
Genetic	100,00%	0,00%	0,00%	0,00%
Incidence by type of cancer	100,00%	0,00%	0,00%	0,00%
Population statistics	100,00%	0,00%	0,00%	0,00%
Prevalence	100,00%	0,00%	0,00%	0,00%
Primary cause of death	100,00%	0,00%	0,00%	0,00%

Tabela 95. Questão P21. Classe: **Therapeutic**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Adverse effect	100,00%	0,00%	0,00%	0,00%
Contraindication	100,00%	0,00%	0,00%	0,00%
Current therapeutic	100,00%	0,00%	0,00%	0,00%
Drug of choice	100,00%	0,00%	0,00%	0,00%
Efficacy	100,00%	0,00%	0,00%	0,00%
Indication	100,00%	0,00%	0,00%	0,00%
Monitoring	100,00%	0,00%	0,00%	0,00%
Name	100,00%	0,00%	0,00%	0,00%
Patient compliance	100,00%	0,00%	0,00%	0,00%

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Prevention	87,50%	12,50%	0,00%	0,00%
Procedure	100,00%	0,00%	0,00%	0,00%
Protocol	100,00%	0,00%	0,00%	0,00%
Type of treatment	100,00%	0,00%	0,00%	0,00%

Tabela 96. Questão 2. Classe: **Pharmacologic Substance**

Propriedade	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Adverse effects	100,00%	0,00%	0,00%	0,00%
Drug	100,00%	0,00%	0,00%	0,00%
Biological effects	100,00%	0,00%	0,00%	0,00%
Brand name	87,50%	0,00%	0,00%	12,50%
Chemical name	100,00%	0,00%	0,00%	0,00%
Dosage	100,00%	0,00%	0,00%	0,00%
Mechanism of action	100,00%	0,00%	0,00%	0,00%
Pharmacalogical action	100,00%	0,00%	0,00%	0,00%
Radioiodine	62,50%	25,00%	0,00%	12,50%
Reaction	100,00%	0,00%	0,00%	0,00%

Questão P23. Foi sugerida a inclusão das classes radioterapia, reabilitação e saúde mental.

Questão P24. Para 87,50%, o modelo de domínio atende aos requisitos para representação de de artigo no domínio da Oncologia, e 12,50% aprovaram com restrições, por esse não incluir as classes radioterapia e reabilitação.

Os resultados dessa ronda foram estruturados em uma planilha (Tabela 97), onde é apresentada a classe e o índice de concordância sobre a associação entre as propriedades e a classe. Nessa ronda, foram consideradas como propriedades a serem mantidas as que obtiveram índice de 100% de aprovação. As demais propriedades que foram avaliadas como "não deve ser mantida como propriedade" ou que houve divergência quanto à associação com a classe foram submetidas a uma segunda ronda.

Tabela 97. **Índice de concordância das propriedades**

Item	Deve ser mantida como propriedade (%)
Classe: Publisher	
Name	100,00
Homepage	87,50
Location	87,50

Item	Deve ser mantida como propriedade (%)
Date created	87,50
Classe: Journal	
Title	100,00
Abbreviation	87,50
Identifier	100,00
Publisher	100,00
Date Created	100,00
Homepage	100,00
Score	100,00
Classe: Scientific Article	
Title	100,00
Identifier	100,00
Volume	100,00
Issue	100,00
Page range	87,50
Date Issued	100,00
Author	100,00
Abstract	100,00
Subject	87,50
Language	87,50
References	87,50
Appendix	62,50
Illustration	75,00
Type	75,00
Format	87,50
File size	62,50
Duration	50,00
Rights	75,00
Access Mode	87,50
Score	87,50
Funding Agency	62,50
GrantNumber	62,50
Level of Evidence	100,00
Conflit of Interest	75,00
Relation	75,00
Status	75,00
Classe: Author	
Name	100,00
Identifier	87,50
Affiliation	87,50
Mailbox	100,00
Score	87,50
Classe: Institution	

Item	Deve ser mantida como propriedade (%)
Name	100,00
Homepage	87,50
Location	87,50
Type	87,50
Classe: Funding Agency	
Name	75,00
Homepage	75,00
Location	75,00
Type	75,00
Classe: Illustration	
Caption	75,00
Description	62,50
Type	75,00
Classe: Note	
Author	62,50
Date created	62,50
Mailbox	62,50
Relation	75,00
Description	62,50
Classe: Gene	
Expression level	87,50
Gene id	87,50
Gene symbol	87,50
Classe: **Organism**	
Age	87,50
Ethnicity	100,00
Gender	100,00
Race	87,50
Description	100,00
Ethnicity	100,00
Classe: Design Research	
Data collection technique	87,50
Research method	100,00
Statistical technique	100,00
Geographic location	87,50
Sample	75,00
Sample size	75,00
Study Subject	100,00
Classe: Anatomy	
Anatomic region	100,00
Biological process	87,50
Cell componente	87,50
Propriedade	**Deve ser mantida como propriedade (%)**

Classe: Anatomy	
Tissue type	100,00
Description	87,50
Classe: Neoplasm	
Cause	87,50
Classification	100,00
Date of staging	87,50
Grade	87,50
Lesion characteristic	100,00
Biological mechanism	100,00
Metastasis	100,00
Recurrence	100,00
Site	87,50
Size	87,50
Status at staging	100,00
Type of cancer	100,00
Classe: Disease	
Cause	100,00
Characteristic	100,00
Class of disease	100,00
Complication	100,00
Description	87,50
Disease agente	100,00
Biological mechanism	100,00
Prognosis	100,00
Rare disease supplementary	87,50
Risk factors	100,00
Classe: Diagnosis	
Category	87,50
Clinically condition	87,50
Condition/ finding	87,50
Interpretation of clinical finding	87,50
Criteria/ manifestations	87,50
Pathology findings	100,00
Previous diagnosis	87,50
Sign or symptom	75,00
Vital status	87,50
Classe: Diagnostic procedure	
Biomarker	87,50
Chemical agent administration	100,00
Diagnostic criteria	87,50
Staging system	87,50
Technique	87,50
Test	100,00
Propriedade	**Deve ser mantida como propriedade (%)**

Classe: Diagnostic procedure	
Type of procedure	100,00
Classe: Imagiology	
Anatomic frame of reference	87,50
Caption	87,50
Color depth	62,50
Date created	87,50
Description	87,50
Equipment	75,00
Format	75,00
Magnification	87,50
Physical description	75,00
Primary magnification	50,00
Processing	75,00
Provenance	87,50
Size of pixel	62,50
Software	75,00
Classe: Epidemiology	
Causality/ association	100,00
Clinical trial information	87,50
Etiology	100,00
Genetic	100,00
Incidence by type of câncer	100,00
Population statistics	100,00
Prevalence	100,00
Primary cause of death	100,00
Classe: Therapeutic	
Adverse effect	100,00
Contraindication	100,00
Current therapeutic	100,00
Drug of choice	100,00
Efficacy	100,00
Indication	100,00
Monitoring	100,00
Name	100,00
Patient compliance	100,00
Prevention	87,50
Procedure	100,00
Protocol	100,00
Type of treatment	100,00
Classe: Pharmacologic Substance	
Adverse effects	100,00
Drug	100,00
Biological effects	100,00
Propriedade	**Deve ser mantida como propriedade (%)**

Classe: Pharmacologic Substance	
Brand name	87,50
Chemical name	100,00
Dosage	100,00
Mechanism of action	100,00
Pharmacalogical action	100,00
Radioiodine	62,50
Reaction	100,00

8.1.2 Estudo Delphi – segunda ronda

Amostra

Com base nesses resultados, foi formulado um novo questionário submetido aos oito bibliotecários que participaram da primeira ronda.

Questionário

Primeiramente, os dados foram agrupados, analisados e organizados por percentuais obtidos na primeira ronda. Para tanto, o questionário foi estruturado em 23 páginas e 31 perguntas, de modo que cada página contivesse uma classe e questões relacionadas com as propriedades (APÊNDICE E).

Como o objetivo da técnica Delphi é obter consenso entre os participantes, o questionário, na segunda ronda, apresentou o percentual de respostas obtidas para todas propriedades em relação às questões formuladas na primeira ronda: manutenção das propriedades nas classes correspondentes; não manutenção da propriedade; e associação da propriedade com outra classe.

Para as propriedades em que houve divergências, as definições das propriedades e as classes foram reformuladas para melhor clarificação. A cada questão foi fornecido o percentual obtido quanto à manutenção da propriedade e solicitado que o participante indicasse se concordava fortemente, concordava ou não concordava. Os respondentes foram informados de que, para manutenção da propriedade no modelo, ela deveria obter um percentual de concordância igual ou superior a 75%, conforme preconizado por Murry & Hammons (1995). Portanto, para que uma propriedade alcançasse o consenso, deveria atingir o percentual de 75%.

A partir da sugestão de um dos participantes na primeira ronda, foi acrescida a propriedade "acompanhamento psicológico". Essa propriedade também foi submetida à avaliação de todos os respondentes. Foram feitos comentários que, quando pertinentes, foram incluídos; outros não o foram

por já estarem contemplados pelas propriedades existentes. Alguns questionamentos precisaram ser esclarecidos por *e-mail*.

Na segunda ronda, foi utilizada a escala tipo Likert com três pontos, composta por: "Concordo fortemente", que equivale ao maior nível de concordância; "Concordo", que corresponde a uma percepção neutra sobre a propriedade; e "Não concordo", que representa o nível de não concordância. O número reduzido de pontos (3) foi priorizado com intuito de minimizar o desgaste dos respondentes, visto que se trata de um questionário longo aplicado pela segunda vez. Identificou-se na literatura que a escala com esse número de pontos se ajusta a pequenas amostras, demanda pouco tempo de resposta e produz opções de respostas suficientes (Dalmoro & Vieira, 2014, Rodriguez, 2005)

O questionário aplicado na segunda ronda foi estruturado de modo que os respondentes pudessem visualizar os resultados (índices) obtidos na primeira ronda.

Cada classe foi apresentada em página distinta do questionário. As propriedades que obtiveram consenso (100%) quanto a sua manutenção, na primeira ronda, foram listadas separadamente sem necessidade de avaliação pelos respondentes (Tabela 98). As propriedades que apresentaram divergência quanto a sua manutenção, ou a sua vinculação à classe que representava, foram agrupadas de acordo com o índice de avaliação obtido na primeira ronda (Tabela 99):

a) 88% dos respondentes aprovaram a permanência da propriedade;

b) 75% dos respondentes aprovaram a permanência da propriedade; e

c) igual ou inferior a 63% dos respondentes optaram por manter a propriedade.

Resultado

A Tabela 98 apresenta as classes com as propriedades que obtiveram 100% de aprovação na primeira ronda e, que, ao serem incluídas no questionário da segunda ronda, não receberam nenhum comentário.

Tabela 98. **Propriedades com 100% de aprovação na primeira ronda.**

Classe/Propriedade	Índice de aprovação na primeira I ronda
Classe: Publisher	
Name	100,00
Classe: Journal	
Title	100,00
Identifier	100,00
Publisher	100,00
Date Created	100,00
Homepage	100,00
Score	100,00
Classe: Scientific Article	
Title	100,00
Identifier	100,00
Volume	100,00
Issue	100,00
Date Issued	100,00
Author	100,00
Abstract	100,00
Level of Evidence	100,00
Classe: Author	
Name	100,00
Mailbox	100,00
Classe: Institution	
Name	100,00
Classe: **Organism**	
Ethnicity	100,00
Gender	100,00
Description	100,00
Ethnicity	100,00
Classe: Design Research	
Research method	100,00
Statistical technique	100,00
Study Subject	100,00
Classe: Anatomy	
Anatomic region	100,00
Tissue type	100,00
Classe: Neoplasm	
Classification	100,00
Lesion characteristic	100,00
Biological mechanism	100,00
Metastasis	100,00
Recurrence	100,00
Status at staging	100,00
Type of cancer	100,00

Classe: Disease	
Cause	100,00
Characteristic	100,00
Class of disease	100,00
Complication	100,00
Disease agente	100,00
Biological mechanism	100,00
Prognosis	100,00
Risk factors	100,00
Classe: Diagnosis	
Pathology findings	100,00
Classe: Diagnostic procedure	
Chemical agent administration	100,00
Test	100,00
Type of procedure	100,00
Classe: Epidemiology	
Causality/ association	100,00
Etiology	100,00
Genetic	100,00
Incidence by type of câncer	100,00
Population statistics	100,00
Prevalence	100,00
Primary cause of death	100,00
Classe: Therapeutic	
Adverse effect	100,00
Contraindication	100,00
Current therapeutic	100,00
Drug of choice	100,00
Efficacy	100,00
Indication	100,00
Monitoring	100,00
Name	100,00
Patient compliance	100,00
Procedure	100,00
Protocol	100,00
Type of treatment	100,00
Classe: Pharmacologic Substance	
Adverse effects	100,00
Drug	100,00
Biological effects	100,00
Chemical name	100,00
Dosage	100,00
Mechanism of action	100,00
Pharmacalogical action	100,00
Reaction	100,00

A Tabela 99 apresenta as propriedades, os índices de aprovação obtidos na primeira ronda e o resultado da segunda ronda. Na segunda ronda percebe-se que não houve grandes mudanças quanto a permanência da propriedade (soma "Concordo fortemente" e "Concordo" superior 75%), mas sim o nível de concordância, o que não indica a sua eliminação.

Tabela 99. **Propriedades submetidas à validação na II ronda**

Questão	Resultado I ronda	Validação II ronda		
	Manutenção da propriedade	Concordo fortemente	Concordo	Não concordo
P1 Class: Publisher				
Homepage	88%	75,00	25,00	0,00
Location	88%	71,43	14,29	14,29
Date created	88%	42,86	42,86	14,29
P2 Class: Journal				
Abbreviation	88%	75,00	25,00	0,00
P3 Class: Scientific Article				
Page range	88%	75,00	25,00	0,00
Subject	88%	87,50	12,50	0,00
Language	88%	62,50	37,50	0,00
References	88%	87,50	12,50	0,00
Format	88%	62,50	37,50	0,00
Access Mode	88%	75,00	25,00	0,00
Score	88%	62,50	37,50	0,00
P4 Class: Scientific Article				
Rights	75%	25,00	75,00	0,00
Conflit of Interest	75%	37,50	62,50	0,00
Relation	75%	75,00	25,00	0,00
Status	75%	37,50	62,50	0,00
Illustration	75%	50,00	50,00	0,00
Type	75%	75,00	25,00	0,00
P5 Class: Scientific Article				
Appendix	63%	25,00	75,00	0,00
Funding Agency	63%	25,00	75,00	0,00
GrantNumber	63%	25,00	25,00	50,00
Duration	63%	62,50	25,00	12,50
File size	63%	50,00	50,00	0,00
P6 Class: Author				
Identifier	88%	62,50	25,00	12,50
Affiliation	88%	62,50	37,50	0,00
Score	88%	62,50	25,00	12,50
P7 Class: Institution				
Homepage	88%	50,00	50,00	0,00
Location	88%	50,00	37,50	12,50
Type	88%	50,00	50,00	0,00
P8 Class: Funding Agency				
Homepage	75%	25,00	75,00	0,00
Location	75%	25,00	62,50	12,50
Type	75%	37,50	62,50	0,00

Questão	Resultado I ronda	Validação II ronda		
	Manutenção da propriedade	Concordo fortemente	Concordo	Não concordo
P9 Class: Design Research				
Data collection technique	88%	75,00	25,00	0,00
Geographic location	88%	75,00	25,00	0,00
P10 Class: Design Research				
Sample	75%	75,00	25,00	0,00
Sample size	75%	75,00	25,00	0,00
P11 Class: Illustration				
Caption	75%	37,50	25,00	0,00
Type	75%	50,00	25,00	0,00
P12 Class: Illustration				
Description	63%	37,50	37,50	25,00
P13 Class: Note				
Relation	75%	50,00	37,50	12,50
P14 Class: Note				
Author	63%	37,50	37,50	25,00
Date created	63%	37,50	50,00	12,50
Mailbox	63%	37,50	50,00	12,50
Description	63%	57,14	28,57	14,29
P15 Class: Gene				
Expression level	88%	75,00	12,50	12,50
Gene id	88%	75,00	12,50	12,50
Gene symbol	88%	75,00	12,50	12,50
P16 Class: Organism				
Age	88%	75,00	12,50	12,50
Race	88%	62,50	25,00	12,50
P17 Class: Anatomy				
Biological process	88%	62,50	37,50	12,50
Cell component	88%	62,50	37,50	12,50
Protein structure	88%	62,50	37,50	0,00
Description	88%	62,50	37,50	0,00
P18 Class: Neoplasm				
Cause	88%	75,00	25,00	0,00
Date of staging	88%	62,50	37,50	0,00
Grade	88%	62,50	37,50	0,00
Site	88%	62,50	37,50	0,00
Size	88%	75,00	25,00	0,00
P19 Class: Disease				
Description	88%	87,50	12,50	0,00
Rare disease supplementary	88%	75,00	25,00	0,00
P20 Class: Epidemiology				
Clinical trial information	88%	87,50	12,50	0,00
P21 Class: Diagnosis				
Category	88%	62,50	37,50	0,00
Clinically condition	88%	75,00	25,00	0,00
Condition/ finding	88%	62,50	37,50	0,00

Questão	Resultado I ronda	Validação II ronda		
	Manutenção da propriedade	Concordo fortemente	Concordo	Não concordo
Interpretation of clinical finding	88%	62,50	37,50	0,00
Criteria/ manifestations	88%	62,50	37,50	0,00
Previous diagnosis	88%	62,50	37,50	0,00
Vital status	88%	62,50	37,50	0,00
P22 Class: Diagnosis				
Sign or symptom	75%	62,50	37,50	0,00
P23 Class: Diagnostic procedure				
Biomarker	88%	75,00	25,00	0,00
Diagnostic criteria	88%	62,50	37,50	0,00
Staging system	88%	75,00	25,00	0,00
Technique	88%	75,00	25,00	0,00
P24 Class: Image				
Anatomic frame of reference	88%	37,50	62,50	0,00
Caption	88%	37,50	62,50	0,00
Date created	88%	37,50	62,50	0,00
Description	88%	37,50	62,50	0,00
Magnification	88%	50,00	50,00	0,00
Provenance	88%	50,00	50,00	0,00
Type	88%	37,50	62,50	0,00
Visualization Mode	88%	37,50	50,00	12,50
P25 Class: Image				
Equipment	75%	62,50	37,50	0,00
Format	75%	37,50	50,00	12,50
Physical description	75%	37,50	62,50	0,00
Processing	75%	50,00	50,00	0,00
Software	75%	50,00	50,00	0,00
P26 Class: Image				
Color depth	69%	25,00	62,50	12,50
Primary magnification	69%	25,00	62,50	12,50
Size of pixel	69%	25,00	62,50	12,50
P27 Class: Therapeutic				
Prevention	88%	75,00	25,00	0,00
P28 Class: Therapeutic (inclusão de nova propriedade)				
Psychological support		50,00	50,00	0,00
P29 Class: Pharmacologic Substance				
Brand name	88%	62,50	25,00	12,50
P30 Class: Pharmacologic Substance				
Radioiodine	63%	50,00	37,50	12,50

A Tabela 100 traz as classes e as propriedades que obtiveram o percentual igual ou superior a 75% de aprovação na segunda ronda. As considerações e sugestões apontadas pelos participantes foram analisadas e aquelas avaliadas como procedentes foram incluídas no estudo. Na primeira ronda foi sugerida a inclusão da classe acompanhamento psicológico (*psychological support*) que, embora não

tenha sido acatada como sugestão para classe, foi incluída como propriedade e submetida à validação dos participantes na segunda ronda. Essa propriedade obteve nível de concordância igual a 100%, o que determinou sua inclusão no modelo na classe *Therapeutic*.

A análise comparativa dos resultados da primeira e segunda ronda mostra que houve propriedades que não obtiveram alto índice de concordância, mas atingiram o percentual estabelecido e explicitado para os respondentes para manutenção da propriedade no modelo (75%). A única propriedade que não atingiu esse índice foi a *Grant Number*. No entanto, devido a relevância dessa propriedade para gestão dos projetos sob a tutela de agências de financiamento europeias (Community Research and Development Information Service – CORDIS, Horizon 2020, OpenAire), ela foi mantida.

As propriedades associadas à classe *Note* (*author; date created; mailbox; relation e description*) e *Gene* (*expression level, gene id* e *gene symbol*) obtiveram baixos índices de concordância para sua manutenção, mas ainda assim obtiveram resultado que permite a sua inclusão no modelo.

As propriedades da classe *Imagiology* atingiram nível de concordância de 100% ao somar a indicação "Concordo fortemente" e "Concordo". No entanto, a indicação "Concordo" obteve maior peso, o que pode revelar pouco convencimento quanto à permanência da classe e de suas propriedades.

Como os participantes não tiveram acesso aos vocabulários controlados a serem adotados no perfil de aplicação, acredita-se que isso originou dúvidas em relação a algumas propriedades, por exemplo: a propriedade *classification* na classe *Neoplasm*. Questionou-se se ela compreendia a morfologia e a histologia do câncer. O ICD-03, VES inserido no PAM, possibilita a inclusão de códigos que contemplam os dois aspectos. Dúvidas como essas foram dirimidas por contato telefônico/e-mail.

Tabela 100. **Classes e propriedades validadas pelo Estudo Delphi**.

Classe/Propriedade	Índice de concordância (%)	Índice de discordância (%)
Classe: Publisher		
Name	100,00	0,00
Homepage	100,00	0,00
Location	85,71	14,29
Date created	85,71	14,29
Classe: Journal		
Title	100,00	0,00
Abbreviation	100,00	0,00
Identifier	100,00	0,00
Publisher	100,00	0,00

Classe/Propriedade	Índice de concordância (%)	Índice de discordância (%)
Date Created	100,00	0,00
Homepage	100,00	0,00
Score	100,00	0,00
Classe: Scientific Article		
Title	100,00	0,00
Identifier	100,00	0,00
Volume	100,00	0,00
Issue	100,00	0,00
Page range	100,00	0,00
Date Issued	100,00	0,00
Author	100,00	0,00
Abstract	100,00	0,00
Subject	100,00	0,00
Language	100,00	0,00
References	100,00	0,00
Appendix	100,00	0,00
Illustration	100,00	0,00
Type	100,00	0,00
Format	100,00	0,00
File size	100,00	0,00
Duration	87,50	12,50
Rights	100,00	0,00
Access Mode	100,00	0,00
Score	100,00	0,00
Funding Agency	100,00	0,00
GrantNumber *	50,00	50,00
Level of Evidence	100,00	0,00
Conflit of Interest	100,00	0,00
Relation	100,00	0,00
Status	100,00	0,00
Classe: Author		
Name	100,00	0,00
Identifier	87,50	12,50
Affiliation	100,00	0,00
Mailbox	100,00	0,00
Score	87,50	12,50
Classe: Institution		
Name	100,00	0,00
Homepage	100,00	0,00
Location	87,50	12,50

Classe/Propriedade	Índice de concordância (%)	Índice de discordância (%)
Type	100,00	0,00
Classe: Funding Agency		
Name	100,00	0,00
Homepage	100,00	0,00
Location	87,50	12,50
Type	100,00	0,00
Classe: Illustration		
Caption	100,00	0,00
Description	75,00	25,00
Type	100,00	0,00
Classe: Note		
Author	75,00	25,00
Date created	87,50	12,50
Mailbox	87,50	12,50
Relation	87,50	12,50
Description	85,71	14,29
Classe: Gene		
Expression level	87,50	12,50
Gene id	87,50	12,50
Gene symbol	87,50	12,50
Classe: Organism		
Age	87,50	12,50
Ethnicity	100,00	0,00
Gender	100,00	0,00
Race	87,50	12,50
Description	100,00	0,00
Classe: Design Research		
Data collection technique	100,00	0,00
Research method	100,00	0,00
Statistical technique	100,00	0,00
Geographic location	100,00	0,00
Sample	100,00	0,00
Sample size	100,00	0,00
Study Subject	100,00	0,00
Classe: Anatomy		
Anatomic region	100,00	0,00
Biological process	100,00	0,00
Cell componente	100,00	0,00
Protein structure	100,00	0,00
Tissue type	100,00	0,00

Classe/Propriedade	Índice de concordância (%)	Índice de discordância (%)
Description	100,00	0,00
Classe: Neoplasm		
Cause	100,00	0,00
Classification	100,00	0,00
Date of staging	100,00	0,00
Grade	100,00	0,00
Lesion characteristic	100,00	0,00
Biological mechanism	100,00	0,00
Metastasis	100,00	0,00
Recurrence	100,00	0,00
Site	100,00	0,00
Size	100,00	0,00
Status at staging	100,00	0,00
Type of cancer	100,00	0,00
Classe: Disease		
Cause	100,00	0,00
Characteristic	100,00	0,00
Class of disease	100,00	0,00
Complication	100,00	0,00
Description	100,00	0,00
Disease agent	100,00	0,00
Biological mechanism	100,00	0,00
Prognosis	100,00	0,00
Rare disease supplementary	100,00	0,00
Risk factors	100,00	0,00
Classe: Diagnosis		
Category	100,00	0,00
Clinically condition	100,00	0,00
Condition/ finding	100,00	0,00
Interpretation of clinical finding	100,00	0,00
Criteria/ manifestations	100,00	0,00
Pathology findings	100,00	0,00
Previous diagnosis	100,00	0,00
Sign or symptom	100,00	0,00
Vital status	100,00	0,00
Classe: Diagnosis procedure		
Biomarker	100,00	0,00
Chemical agent administration	100,00	0,00
Diagnostic criteria	100,00	0,00
Staging system	100,00	0,00

Classe/Propriedade	Índice de concordância (%)	Índice de discordância (%)
Technique	100,00	0,00
Test	100,00	0,00
Type of procedure	100,00	0,00
Classe: Image		
Anatomic frame of reference	100,00	0,00
Caption	100,00	0,00
Color depth	87,50	12,50
Date created	100,00	0,00
Description	100,00	0,00
Equipment	100,00	0,00
Format	87,50	12,50
Magnification	100,00	0,00
Physical description	100,00	0,00
Primary magnification	87,50	12,50
Processing	100,00	0,00
Provenance	100,00	0,00
Size of pixel	87,50	12,50
Software	100,00	0,00
Visualization mode	87,50	12,50
Classe: Epidemiology		
Causality/ association	100,00	0,00
Clinical trial information	100,00	0,00
Etiology	100,00	0,00
Genetic	100,00	0,00
Incidence by type of cancer	100,00	0,00
Population statistics	100,00	0,00
Prevalence	100,00	0,00
Primary cause of death	100,00	0,00
Classe: Therapeutic		
Adverse effect	100,00	0,00
Contraindication	100,00	0,00
Current therapeutic	100,00	0,00
Drug of choice	100,00	0,00
Efficacy	100,00	0,00
Indication	100,00	0,00
Monitoring	100,00	0,00
Name	100,00	0,00
Patient compliance	100,00	0,00
Prevention	100,00	0,00
Procedure	100,00	0,00
Protocol	100,00	0,00
Psycological support	100,00	0,00
Type of treatment	100,00	0,00

Classe: Pharmacologic Substance		
Adverse effect	100,00	0,00
Drug	100,00	0,00
Biological effects	100,00	0,00
Brand name	87,50	12,50
Chemical name	100,00	0,00
Dosage	100,00	0,00
Mechanism of action	100,00	0,00
Pharmacalogical action	100,00	0,00
Radioiodine	87,50	12,50
Reaction	100,00	0,00

8.2 *DESCRIPTION SET PROFILE*

O modelo de dados detalhado para artigos científicos no domínio da Oncologia no contexto dos repositórios digitais científicos, apresentado aos oito respondentes para validação (Figura 52), obteve o índice de consenso necessário para sua aprovação e, consequentemente, as classes e propriedades que o compõe. À esse modelo incluíu-se apenas uma nova propriedade, a *psychologicalSupport*, à classe *Therapeutic*.

Na versão final da Matriz de Restrição foi adicionada a propriedade rdf:type. A propriedade rdf:type informa qual é o tipo do recurso, ou seja, informa que o recurso é instância de uma determinada classe. Pode-se utilizar como exemplo a primeira tabela da Matriz de Restrição, que refere à classe Publisher. A a propriedade rdf:type dessa classe é a classe dbpedia-owl:Publisher (Tabelas 101 e 102).

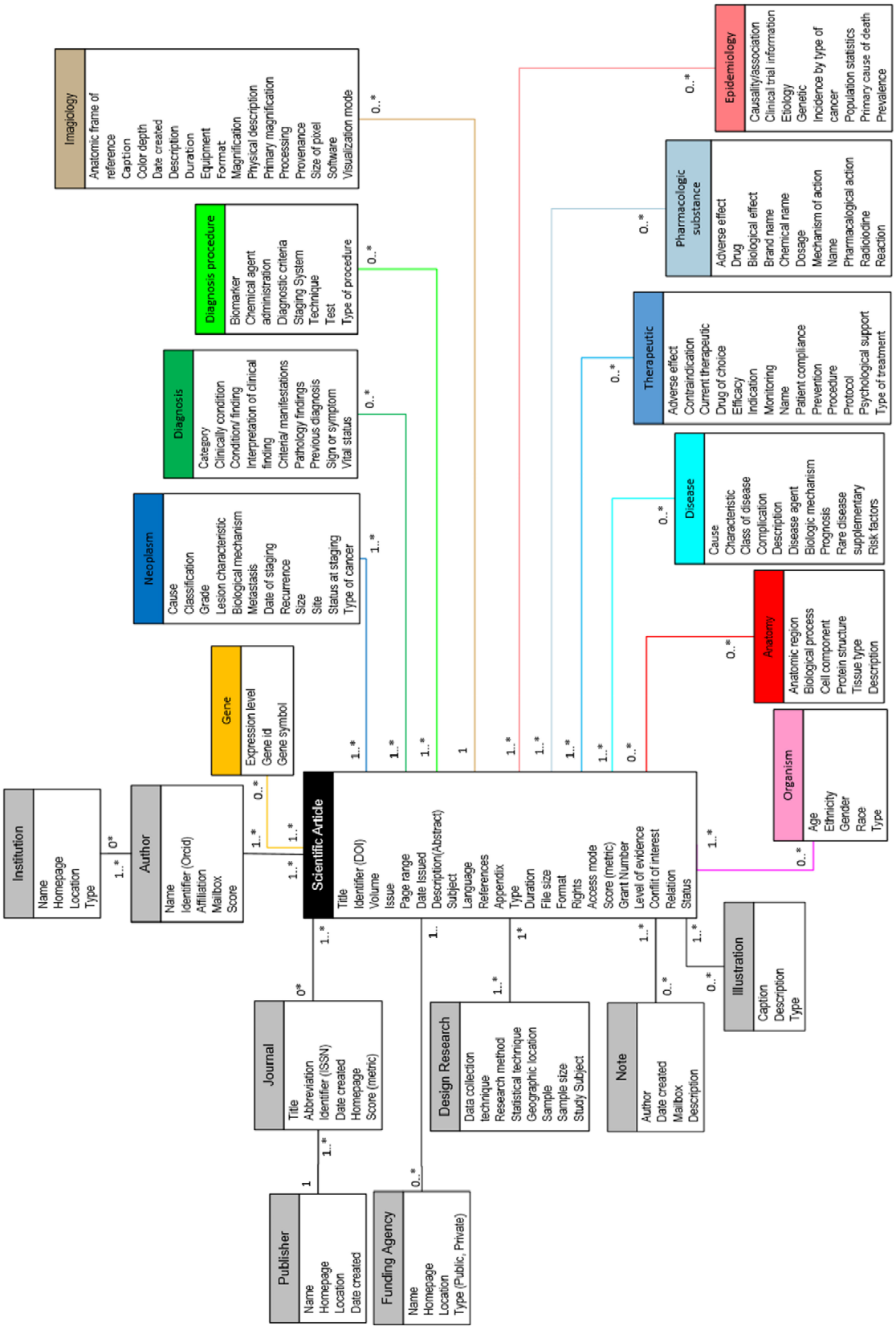


Figura 52. Modelo de dados para artigos científicos no domínio da Oncologia armazenados em repositórios digitais científicos.

Tabela 101. **Definição dos *namespaces* utilizados.**

Título	**Namespace**	**Prefixo**
Bibliographic Ontology Specification	http://purl.org/ontology/bibo/	bibo
The DBPEDIA Ontology	http://dbpedia.org/ontology/	dbpedia-owl
DCMI Metadata Terms	http://purl.org/dc/terms/	dcterms
Friend of a Friend Vocabulary	http://xmlns.com/foaf/0.1/	foaf
Oncology Vocabulary	http://opendata.dsi.uminho.pt/health/	onco
The RDF Schema Vocabulary	http://www.w3.org/2000/01/rdf-schema#	rdfs
Schema.org vocabular	http://schema.org/	schema.org

Tabela 102. **Matriz de restrições**

Description: Publisher				Class: dbpedia-owl:Publisher			Use: Publishing company			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	rdfs:label	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	No	1	URI	Name of entity responsible for making the resource available.
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0 - 1	URI	A homepage for publisher.
Location	dcterms:coverage	Non literal	Yes	------	Yes	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	No	1	URI	The spatial topic of the resource.
Date created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	------	No	0 - 1	Date	Date of formal issuance of the resource.
Type	rdf:type	dbpedia-owl:Publisher	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal.

Description: Journal				Class: bibo:Journal			Use: A periodical of scholarly journal			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Title	dcterms:title	Literal	Yes	------	No	------	No	1	String	A name given to the journal. Ex.: Advances in cancer research.
Abbreviation	dbpedia-owl:abbreviation	Literal	Yes	NLM Catalog (https://www.ncbi.nlm.nih. gov/pubmed/)	No	------	No	1	String	The abbreviated form representing a resource. Ex.: Adv. cancer res.
Identifier	dcterms:identifier	Literal	Yes	https://www.iso.org/standard/39601.html	No	------	No	1	String	An unambiguous reference to the resource. Ex.: ISSN: 0065-230X
Publisher	dcterms:publisher	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	Yes	1..*	URI	An entity responsible for making the resource available. Ex.: Elsevier, http://id.loc.gov/authorities/names/no95025589
Date Created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	------	No	0-1	Date	Date of formal issuance of the resource. Ex.: 1953-03-23.
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0-1	URI	A homepage for resource. Ex.: http://www.sciencedirect.com/science/bookseries/0065230X.
Score (metric)	onco:score	Non literal	No	------	Yes	------	No	0..*	URI	Represents the metric for resource. Ex.: 82 H Index (http://www.scimagojr.com/journalsearch. php?q=28690&tip=sid)
Type	rdf:type	bibo:Journal OU dbpedia-owl:Publisher	No	------	Yes	------	No	1	URI	Publishing company

Description: Scientific Article				Class: schema:ScholarlyArticle			Use: A scholarly academic article, typically published in a journal			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Title	dcterms:title	Literal	Yes	------	No	------	No	1..*	String	A name given to the resource.
Identifier	dc:identifier	Non literal	No	------	Yes	------	No	1..*	URI	An unambiguous reference to the resource within a given context, ex. DOI , Handle
Volume	bibo:volume	Literal	Yes	------	No	------	No	1	String	Volume of a journal.
Issue	bibo:issue	Literal	Yes	------	No	------	No	0-1	String	Issue of a journal.
Page range	bibo:pages	Literal	Yes	------	No	------	No	1	String	Page range of a document.
Date Issued	dcterms:issued	Literal	Yes	http://purl.org/dc/terms/W3CDTF	Yes	------	No	1	Date	Date of publication of the resource.
Author	dcterms:creator.orcid	Non literal	No	------	Yes	https://orcid.org/	Yes	1..*	URI	An entity primarily responsible for making the resource. (Ex.: orcid.org/0000-0002-4055-5882)
Abstract	dcterms:abstract	Literal	Yes	------	No	------	No	1..*	String	A summary of the resource.
Subject	dcterms:subject	Non literal	Yes	------	Yes	http://purl.org/dc/terms/MESH	No	1..*	URI	The topic of the resource.
Language	dcterms:language	Literal	Yes	http://lcweb.loc.gov/standards/iso639-2/langhome.html	No	------	No	1-3	String	A language of the resource. Ex.: eng, por, etc.
References	dc:references	Literal	Yes	------	No	------	No	1	String	A related resource that is referenced, cited, or otherwise pointed to by the described resource.
Appendix	onco:appendix	Non literal	No	------	Yes	------	No	0..*	URI	Text or document prepared by the author
Illustration	onco:illustration	Literal	Yes	------	No	------	Yes	0...*	String	A image, graphic or table within a journal.
Type	dcterms:type	Non literal	No	------	Yes	http://vocabularies.coar-repositories.org/documentation/resource_types/	No	1	URI	The nature or genre of the resource. Ex.: Research, case report, etc.
Format	dcterms:format	Non literal	No	------	Yes	http://www.iana.org/assignments/media-types/	No	1..*	URI	The file format, physical medium, or dimensions of the resource. Ex.: audio, text or video.
Rights	dcterms:rights	Non literal	No	------	Yes	http://creativecommons.org/licenses/by/4.0/	No	1	URI	Information about rights held in and over the resource. Ex.: Licence Creative Commons (CC BY, CC BY-ND, etc.).
Access Mode	onco:accessMode	Non literal	No	------	Yes	http://vocabularies.coar-repositories.org(ongoing implementation)	No	1	URI	Declares the degree of openness of a resource. Ex.: open, embargo or restrict.
Score (metric)	onco:score	Non literal	No	------	Yes	------	No	0..*	URI	Represents the metric being observed. Ex.: Scopus, Scholar Google, Web of Science.
Funding Agency	onco:fundingAgency	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	Yes	1..*	URI	Institutions that offer different forms of financing aimed at conducting research. Ex.: Fundação para a Ciência e a Tecnologia (FTC - Portugal)
GrantNumber	onco:grantNumber	Literal	Yes	------	No	----	No	1..*	String	A data property specifying the grant number of a grant provided by a funding agency. Ex.: UID/Multi/04378/2013
Level of Evidence	schema:evidenceLevel	Non literal	Yes	------	No	http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009/	No	1	String	Strength of evidence of the data used to formulate the guideline. Ex.: Level 1A for treatment: a systematic review of randomized clinical trial.
Conflit of Interest	onco:conflitOfInterest	Literal	Yes	------	No	------	No	0-1	String	A situation in which an individual might benefit personally from oficial or professional actions.
Relation	dcterms:relation	Non literal	No	------	Yes	------	No	0..*	URI	A related resource. Ex.: Thesis, Research project.
Status	bibo:status	Non literal	Yes	------	No	http://www.niso.org/publications/rp/RP-8-2008.pdf	No	1	String	Status of a resource. Ex.: Submitted manuscript under review, Accepted manuscript
File size	dcterms:extend	Literal	Yes	------	No	------	No	0-1	String	File size corresponding to the file of article (ex.: 300MB)
Duration	dcterms:extend	Literal	Yes	------	No	------	No	0-1	String	Duration of the audio or video corresponding to the file of article (ex.: 30 min
Type	rdf:type	schema:ScholarlyArticle	No	------	Yes	------	No	1	URI	A periodical of scholarly journal

Description: Author				Class: dcterms:Agent			Use: A resource that acts or has the power to act			
Label	Property	Range	Valuestring	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	foaf:name	Literal	Yes	------	No	------	No	1	String	An entity primarily responsible for making the resource.
Identifier	dc:identifier	Non literal	Yes	------	Yes	------	No	1	URI	An unambiguous reference to the resource within a given context. Ex.: Orcid
Affiliation	schema:affiliation	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	Yes	0..*	URI	An organization that this person is affiliated.
Mailbox	foaf:mbox	Non literal	No	------	Yes	------	No	0-1	mailto:URI	A personal mailbox.
Score (metric)	onco:score	Non iteral	Yes	------	Yes	------	No	0-1	URI	A numeric values measuring performance or quality of something. Ex.: Google Citation.
Type	rdf:type	dcterms:Agent	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Institution				Class: dbpedia-owl:institution		Use: An organization, establishment or foundation.				
Label	Property	Range	Valuestring	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	rdfs:label	Non literal	Yes	------	Yes	http://id.loc.gov/authorities/names.html	No	1	URI	Name of Institution. Ex.: Universidade do Minho or http://id.loc.gov/rwo/agents/n85096589
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0-1	URI	A homepage for organization.
Location	dcterms:coverage	Non literal	Yes	------	Yes	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	No	1	URI	Physical location of the item.
Type	dcterms:type	Non literal	Yes	------	Yes	https://meshb.nlm.nih.gov/record/ui?ui=D009938	No	0-1	URI	The nature or genre of the resource. Ex.: Academic institution, Hospital, etc.
Type	rdf:type	dbpedia-owl:institution	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Funding Agency				Class: onco:FundingAgency		Use: Insitutions or organizations that support scientific or technologic research				
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Name	rdfs:label	Non literal	Yes	------	No	http://id.loc.gov/authorities/names.html	No	1	URI	Name of agency (Ex.: Fundação para a Ciência e a Tecnologia, http://id.loc.gov/authorities/names/no00038939
Homepage	foaf:homepage	Non literal	No	------	Yes	------	No	0..*	URI	A homepage for organization
Location	dcterms:coverage	Non literal	Yes	------	Yes	http://www.getty.edu/research/tools/vocabularies/tgn/index.html	No	1	URI	Physical location of the institution.
Type	dcterms:type	Non literal	No	------	Yes	https://meshb.nlm.nih.gov/record/ui?ui=D009938	No	1	URI	The nature or genre of the resource.Ex.: public, private
Type	rdf:type	onco:Funding Agency	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Design Research				Class: onco:DesignResearch			Use: Describes how the study performed to answer the research			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Data collection technique	onco:dataCollection	Non literal	Yes	----	Yes	http://www.ddialliance.org/Specification/DDI-CV/ModeOfCollection_1.0.html	No	0..*	URI	Referes to process of gathering data for testing or research purposes.
Research method	onco:researchMethod	Non literal	Yes	----	Yes	http://id.loc.gov/authorities/subjects.html	No	0..*	URI	A method or procedures used in a research.
Statistical technique	onco:statisticalTechnique	Non literal	Yes	----	No	https://stats.oecd.org/glossary/alpha.asp?Let=A	No	0..*	String	A statistical technique used in a research.
Geographic location	dcterms:coverage	Non literal	No	----	Yes	http://purl.org/dc/terms/TGN	No	0..*	URI	Place where the study was developed.
Sample	onco:sample	Literal	Yes	----	No	No	No	1..*	String	The sample of study.
Sample size	onco:sampleSize	Literal	Yes	----	No	No	No	1..*	String	The sample size.
Study Subject	schema:studySubject	Non literal	Yes	----	Yes	https://meshb.nlm.nih.gov/search	No	1..*	URI	A subject investigated by the study. Ex.: Therapy, Drug, Diagnosis.
Type	rdf:type	onco:Design Research	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Illustration				Class: onco:Illustration			Use: A numbered and captioned image, graphic or table within a journal.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Caption	schema:caption	Literal	Yes	----	No	----	No	0..*	String	The caption for this object.
Description	dcterms:description	Literal	Yes	----	No	----	No	0..*	String	An account of the resource.
Type	dc:type	Literal	Yes	----	No	----	No	0..*	String	The nature or genre of the resource. Ex.: graphic, table or figure.
Type	rdf:type	onco:Illustration	No	------	Yes	----	No	1	URI	A scholarly academic article, typically published in a journal

Description: Note				Class: bibo:Note		Use: A note about a Resource				
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Author	dc:creator	Literal	Yes	----	No	----	Yes	1	String	Na entity primarily responsible for making the resource. (Ex.: Smith, J.)
Date created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	1	Date	Date of creation of the resource
Mailbox	foaf:mbox	Non literal	No	----	Yes	----	No	0-1	mailto: URI	A personal mailbox
Relation	dcterms:relation	Non literal	No	----	Yes	----	No	1	URI	A related resource.
Description	dcterms:description	Literal	Yes	----	No	----	No	1	String	An account of the resource.
Type	rdf:type	bibo:Note	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Gene				Class: dbpedia-owl:Gene			Use: The functional and physical unit of heredity passed from parente to offspring.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Expression level	onco:proteinExpression	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/projects/CCDS/CcdsBrowse.cgi	No	0-1	URI	The determination of the level of gene transcription in a biological specimen.
Gene id	onco:geneIdentifier	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/gene	No	1	URI	A unique gene name within a specific repository, database, or collection.
Gene symbol	onco:geneSymbol	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/gene	No	1	URI	A unique gene name approved by an organism specific nomenclature committee.
Type	rdf:type	dbpedia-owl:Gene	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Organism				Class: onco:Organism			Use: A living entity. Ex.: Animal, human being			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Age	dbpedia-owl:age	Literal	Yes	----	No	----	No	0-1	String	Elapsed time since birth.
Ethnicity	dbpedia-owl:ethnicity	Non literal	Yes	----	No	https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html	No	0-1	String	Members share a common history and a sense of identification with the group
Gender	foaf:gender	Non literal	No	----	Yes	http://hl7.org/fhir/v3/AdministrativeGender	No	0-1	URI	The gender of this agent. Ex.: male or female.
Race	dbpedia-owl:race	Non literal	Yes	----	No	https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html	No	0-1	String	An arbitrary classification of a taxonomic group that is a division of a species.
Description	dcterms:description	Non literal	No	----	Yes	https://meshb.nlm.nih.gov/search	No	1..*	URI	Description of a living entity. Ex.: Animal or human being.
Type	rdf:type	onco:Organism	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Anatomy				Class: onco:Anatomy			Use: Physical structure of animals and plants			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Anatomic region	onco:anatomicRegion	Non literal	No	----	Yes	https://meshb.nlm.nih.gov/search	No	1..*	URI	Identifies areas of the body.
Biological process	onco:biologicalProcess	Lliteral	Yes	----	No	----	No	0..*	String	An activity occurring within an organism, between organisms or among organisms and the mechanisms underlying such events.
Cell component	onco:cellComponent	Non literal	No	----	Yes	https://meshb.nlm.nih.gov/search	No	0..*	URI	The components of the cell.
Protein structure	onco:proteinStructure	Non literal	No	----	Yes	https://www.ncbi.nlm.nih.gov/structure/	No	0..*	URI	The amino acids and their 3-D arrangement in constituting a protein.
Tissue type	onco:tissueType	Non literal	No	----	Yes	https://meshb.nlm.nih.gov/search	No	0..*	URI	Describes collections of differentiated cells, such as epithelium; connective tissue; muscles; nerve tissue.
Description	dcterms:description	Literal	Yes	----	No	----	No	1	String	Describes the referenced anatomy.
Type	rdf:type	onco:Anatomy	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Neoplasm				Class: onco:Neoplasm			Use: A benigno or malignant tissue growth resulting from uncontrolled cell proliferation			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Cause	schema:cause	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence.
Classification	onco:classification	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	1	URI	The classification of cancers by histological type; primary site; location where it was first developed.
Date of staging	onco:dateOfStage	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	0-1	Date	Date of he process of defining at what point in the natural history of a malignant disease a patient is when a diagnosis is made.
Grade	onco:gradeTumor	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0-1	URI	A position on a scale of intensity or amount or quality.
Lesion characteristic	onco:lesionCharacteristic	Literal	Yes	----	No	----	No	0..*	String	Characteristic of localized pathological or traumatic structural change, damage, deformity, or discontinuity of tissue, organ, or body part.
Biological mechanism	schema:mechanismBiologic	Non literal	Yes	----	Yes	http://apps.who.int/classifications/icd10/browse/2016/en	No	0..*	URI	Mechanism biologic associated with disease.
Metastasis	onco:metastasis	Non literal	Yes	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0..*	URI	The spread or migration of cancer cells from one part of the body to another.
Recurrence	onco:recurrence	Literal	Yes	----	No	----	No	0-1	String	If the cancer returned.
Site	onco:site	Non literal	No	----	Yes	http://codes.iarc.fr/usingicdo.php	No	2	URI	A position in which cancer develops (histologic and morphologic).
Size	onco:size	Non literal	Yes	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0-1	URI	The physical magnitude of tumor.
Status at staging	onco:statusAtStaging	Non literal	Yes	----	Yes	http://codes.iarc.fr/usingicdo.php	No	0-1	URI	Level of extent of the cancer within the body.
Type of cancer	onco:typeOfCancer	Non literal	No	----	Yes	https://meshb.nlm.nih.gov/search	No	1	URI	Type of cancer.
Type	rdf:type	onco:Neoplasm	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Disease				Class: dbpedia-owl:Disease			Use: A health problem with certain characteristics or symptoms			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Cause	schema:cause	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence.
Characteristic	onco:characteristic	Literal	Yes	----	No	----	No	1..*	String	The distinguishing of a disease aspect.
Class of disease	onco:classOfDisease	Non literal	No	----	Yes	http://www.icd10data.com/	No	1..*	URI	Classification of diseases. Ex.: Breast cancer.
Complication	onco:complication	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	Any disease or disorder that occurs during the course of, or because of, another disease, treatment, or procedure.
Description	dc:description	Literal	Yes	----	No	----	No	1	String	Description of the disease.
Disease agente	onco:diseaseAgent	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	Any vírus, micro-organism, or other substance causing disease.
Mechanism biologic	schema:mechanismBiologic	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	Mechanism biologic associated with disease.
Prognosis	onco:prognosis	Literal	Yes	----	No	----	No	0..*	String	Prognosis.
Rare disease supplementary	onco:rareDisease	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	Rare disease identification.
Risk factors	schema:riskFactor	Non literal	Yes	----	Yes	http://www.icd10data.com/	No	0..*	URI	A modifiable or non-modifiable fator that increases the risk of a person contracting a condition. Ex.: Tobacco.
Type	rdf:type	dbpedia-owl:Disease	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Diagnosis				Class: onco:Diagnosis			Use: The investigation, analysis and recognition of the presence and nature of disease, condition, or injury from expressed signs and symptoms.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Category	schema:category	Non literal	No	----	Yes	http://www.icd10data.com/	No	1	URI	A category for the item. Ex.: Category based on image clinical report – BI-RADS.
Clinically condition	schema:healthCondition	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	Specifying the health condition(s) of a patient, medical study, or other target audience.
Condition/ finding	onco:diseaseEvidence	Non literal	No	----	Yes	https://loinc.org/	No	0..*	URI	Clinical, laboratory or molecular evidence, or lack of evidence of the disease.
Interpretation of clinical finding	onco:interpretingClinicalFinding	Non literal	No	----	Yes	https://loinc.org/	No	0..*	URI	Interpretation of clinical finding.
Criteria/ manifestations	onco:criteriaManifestation	Literal	Yes	----	No	----	No	0-1	String	Criteria based on patient symptoms and signs, test results. Ex.: Amsterdam criteria.
Pathology findings	onco:pathologyFindings	Non literal	No	----	Yes	https://loinc.org/	No	0..*	URI	Pathology findings.
Previous diagnosis	onco: previous Diagnosis	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	Previous diagnosis.
Sign or symptom	schema:signOrSymptom	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	A sign o r symptom of this condition.
Vital status	onco:vitalStatus	Literal	Yes	----	No	----	No	0..*	String	The state or condition of being living or deceased; also includes the case where the vital status is unknown.
Type	rdf:type	onco:Diagnosis	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Diagnosis procedure				Class:schema: DiagnosticProcedure			Use: A medical procedure intented primarily for diagnostic, as opposed to therapeutic, purposes.			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Biomarker	onco:biomaker	Non literal	Yes	----	Yes	https://loinc.org/	No	0..*	URI	Biomarker used.
Chemical agent administration	onco:chemicalAgentAdministration	Non literal	Yes	----	Yes	https://meshb.nlm.nih.gov/search	No	0..*	URI	Administration of chemical agent in exams.
Diagnostic criteria	onco:diagnosticCriteria	Literal	Yes	----	No	No	No	0..*	String	A standard from which a judgment concerning a medical diagnosis can be established.
Staging System	onco:stagingSystem	Non literal	No		Yes	http://www.icd10data.com/	No	0..*	URI	A system adoped to describe the extent of cancer in a patient's body. Ex.: TNM Staging System.
Technique	onco:technique	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	Yes	1-..*	URI	Encompass all investigations and tests intended to identify the cause of disease. Ex.: laboratory tests and imaging techniques, Differential diagnosis.
Test	onco:test	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	0..*	URI	Process of determining the disease state of a person. Ex.: cancer mutation test, blood chemistry test, etc.
Type of procedure	schema:procedureType	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1-..*	URI	The type of procedure, for example surgical, noninvasive or percutaneous. Ex.: Clinical diagnosis; computer-aided diagnosis.
Type	rdf:type	schema: DiagnosticProcedure	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Imagiology				Class: schema:ImageTest			Use: Any medical imaging modality typically used for diagnostic purposes			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Anatomic frame of reference	onco:anatomicFrameOfReference	Literal	Yes	----	No	----	No	0..*	String	Image acquisition module.
Caption	schema:caption	Literal	Yes	----	No	----	No	1	String	The caption for this object.
Color depth	onco:colorDepth	Literal	Yes	----	No	----	No	0-1	String	Positive integer indicating the desired minimum total number of bits per color channel used to present a pixel.
Date created	dcterms:created	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	----	No	1	Date	Date of creation of the resource.
Description	dcterms:description	Literal	Yes	----	No	----	No	0..*	String	An account of the image.
Equipment	dbpedia-owl:equipment	Literal	Yes	----	No	----	No	0..*	String	Equipment used to analyze, enhance, interpret or display digital images
Format	dc:format	Literal	Yes	http://purl.org/dc/terms/W3CDTF	No	http://www.iana.org/assignments/media-types/	No	0-1	String	The duration of the item. Ex.: For a video, T19:20+01:00.
Magnification	onco:magnification	Literal	Yes	----	No	----	No	0-1	String	Magnification is related to scaling up visuals or images to be able to see more detail.
Physical description	onco:physicalDescription	Literal	Yes	----	No	----	No	0-1	String	Physical description of the image.
Primary magnification	onco:primaryMagnification	Literal	Yes	----	No	----	No	0-1	String	Initial magnification.
Processing	onco:processingImage	Literal	Yes	----	No	----	No	0...*	String	The use of data processing systems to analyze, enhance, interpret or display digital images.
Provenance	dcterms:provenance	Non literal	Yes	----	No	----	No	1	URI	Ownership and custody of the resource of since its creation. Ex.: Hospital, Laboratory.
Size of pixel	onco:sizeOfPixel	Literal	Yes	----	No	----	No	0-1	String	Image size expressed in pixels.
Software	onco:software	Literal	Yes	----	No	----	No	0-1	String	A computer program in source of compiled form.
Type	dcterms:type	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1*	URI	The nature or genre of the resource. Ex.: radiography, videolaryngoscopy.
Type	rdf:type	schema:Imaging Test	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal.
Visualization Mode	onco:visualizationMode	Literal	Yes		No	----	No	0-1	String	Options for visualization (ex.: 3D image).

Description: Therapeutic				Class: onco:Therapeutic			Use: Relating to or involved in therapy			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Adverse effect	onco:adverseEffect	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	An unexpected medical problem that happens during treatment with a drug or other therapy.
Contraindication	schema:contraindication	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	A contraindication for this therapy
Current therapeutic	onco:currentTherapeutic	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	0..*	URI	The name of the therapy that is currently being administered.
Drug of choice	schema:drugClass	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	The class of drug this belongs to (ex., statins)
Efficacy	onco:efficacy	Literal	Yes	----	No	----	No	0..*	String	The ability of a chemical agent or a device to treat a disease.
Indication	schema:indication	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	0..*	URI	A fator that indicates use of this therapy for treatment and/or prevention of a condition, symptom, etc.
Monitoring	onco:monitoring	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	0..*	URI	How a person or a condition is assited or verified.
Patient compliance	onco:patientCompliance	Literal	Yes	----	No	----	No	0..*	String	The act of following a medical regimen or schedule correctly and consistently, including taking medicines or following a diet.
Prevention	onco:prevention	Literal	Yes	----	No	----	No	0..*	String	Describes methods of preventing cancer including limit of exposure to carcinogens and infectious agents, diet, education, etc
Procedure	onco:medicalProcedure	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1..*	URI	Describes a treatment or action taken to prevent or treat disease
Protocol	onco:protocol	Literal	Yes	----	No	----	No	0..*	String	Describes a rule which guides how an activity should be performed.
Type of treatment	onco:typeOfTreatment	Non literal	No	----	Yes	http://www.icd10data.com/ICD10PCS	No	1..*	URI	Cancer treatment adopted to non directed or directed cancer Ex.: control, remove or destroy cancer or prolong a patient's live, alleviate pain.
Type	rdf:type	onco:Therapeutic	No	-----	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal
Psychological assessment	onco:psychologicalAssessment	Literal	Yes	-----	No	------	No	0..*	String	Represents information regarding the psychological assistance to the patient

Description: Pharmacologic Substance				Class: onco:PharmacologicSubstance			Use: Any substance, other than food, that is used to prevent, diagnose, treat or relieve symptoms of a disease or abnormal condition			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Adverse effects	onco:adverseEffect	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	An unexpected medical problem that happens during treatment with a drug or other therapy.
Drug	onco:drug	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	Specifying a drug or medicine used in a medication procedure
Biological effects	onco:biologicalEffect	Non literal	Yes	----	No	http://www.icd10data.com/	No	0..*	URI	Biological effect that the any product provide.Ex.: Radiation
Brand name	dbpedia-owl:brand	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	Brand name. Ex.: Ibuprofen 400 to Ibuprofen.
Chemical name	onco:chemicalName	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	1..*	URI	Chemical name. Ex.: Ibuprofen; Brufen to Ibuprofen

Description: Pharmacologic Substance				Class: onco:PharmacologicSubstance			Use: Any substance, other than food, that is used to prevent, diagnose, treat or relieve symptoms of a disease or abnormal condition			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Dosage	schema:doseValue	Literal	Yes	----	No	----	No	0..*	String	The value of the dose. Ex. 500 mg
Mechanism of action	schema:mechanismOfAction	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	0..*	URI	The specific biochemical interaction through which this drug or supplement produces its pharmacological effect
Pharmacalogical action	onco:pharmacologicAction	Non literal	No	----	Yes	https://mor.nlm.nih.gov/RxNav/	No	0..*	URI	Actions of drugs and their effects on living organisms
Radioiodine	onco:radioiodine	Literal	Yes	----	No	----	No	0..*	String	Use of radioiodine.
Reaction	onco:reaction	Non literal	No	----	Yes	http://www.icd10data.com/	No	0..*	URI	A description of the action in response to some kind of stimulus.
Type	rdf:type	ScholarlyArticle	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

Description: Epidemiology				Class: onco:Epidemiology			Use: The study of the patterns, causes, and control of disease in groups of people			
Label	Property	Range	Value string	SES URI	Value URI	VES URI	Related description	Cardinality	Type	Usage
Causality/ association	onco:causality	Literal	Yes	----	No	----	No	0..*	String	Relationship exists between a cause or effect. Ex.: radiation.
Clinical trial information	onco:clinicalTrialInformation	Literal	Yes	---=p	No	----	No	0-1	String	Report of a pre-planned medical or veterinary clinical case study involving a randomized, evidence-controlled clinical trial.
Etiology	onco:etiology	Non literal	Yes	----	No	http://www.icd10data.com/	No	0..*	URI	The cause of a disease or abnormal condition.
Genetic	onco:geneticEpidemiology	Literal	Yes	----	No	----	No	0-1	String	Interdisciplinary clinical, epidemiologic, genetic, and laboratory studies of persons, families, and populations at high risk of cancer.
Incidence by type of câncer	onco:incidence	Literal	Yes	----	No	----	No	1..*	String	The relative frequency of occurrence of cancer.
Population statistics	onco:populationStatistics	Literal	Yes	----	No	----	No	0..*	String	A value, which represents a particular population characteristic.
Prevalence	onco:prevalence	Literal	Yes	----	No	----	No	0..*	String	The total number of cases of a given disease in a specified population at a designated time.
Primary cause of death	onco:primaryCauseofDeath	Non literal	Yes	----	No	http://www.icd10data.com/	No	0-1	URI	The first significant event which ultimately led to death.
Type	rdf:type	onco:Epidemiology	No	------	Yes	------	No	1	URI	A scholarly academic article, typically published in a journal

SÍNTESE DA SECÇÃO 8

A validação do modelo de dados, que contribuiu para a definição do *Description Set Profile*, foi realizada por profissionais da informação que atuam no domínio da Oncologia. São profissionais que, no exercício da sua profissão, executam atividades associadas aos sistemas de informação, seja na organização e na estruturação, seja no tratamento e na recuperação da informação. Essa atuação possibilita que esses profissionais possam validar o artefacto que foi desenvolvido a partir de demandas (requisitos) de informações de médicos e investigadores ao buscarem por artigos científicos.

Os bibliotecários(as) validaram o modelo de dados (classes, propriedades e cardinalidade) apresentado por meio da Técnica Delphi, aplicada em duas rondas. O estudo Delphi foi realizado com oito profissionais que, além de se posicionarem em relação ao modelo apresentado, contribuíram com sugestões. Todas as classes e propriedades foram mantidas, com adição de uma nova propriedade (*psychologicalSupport*) à classe *Therapeutic.*

Como resultado final do processo de validação foi apresentado o *Description Set Profile*.

9 CONSIDERAÇÕES FINAIS

A mente que se abre a uma nova ideia jamais volta ao seu tamanho original.

Albert Einstein

9.1 INTRODUÇÃO

O aumento e a diversidade de informações não são problemas exclusivos dos dias atuais. Investigadores vêm demonstrando preocupações em relação a esses fatos desde a década de 1960. No entanto, a rapidez e o volume com que essas informações estão sendo produzidas e compartilhadas na Web suscitaram novas inquietações, nomeadamente em relação à produção científica. São questões associadas ao sistema de revisão por pares, ao tipo de acesso, aos direitos autorais, e, sobretudo, referente à descoberta dos recursos. A Web requer ferramentas e métodos para prover os utilizadores com informações relevantes em menos tempo e com menos esforço. A semântica associada à adoção das diretrizes *Linked Data* é vista como alternativa para otimização da descoberta dos recursos, ao proporcionar uma estrutura em que a informação possui um significado que pode ser entendido por agentes inteligentes e, ao mesmo tempo, estabelecer conexões entre outras informações e dados (Web Semântica).

À medida que diretrizes são definidas para alcançar esse modelo, iniciativas como a identificação e a utilização dos tipos de necessidades de informação dos utilizadores em domínios específicos também podem maximizar a descoberta dos recursos. Nesse sentido, podem ser consideradas diferentes granularidades de descrição dos conteúdos e, assim, valorizar o conhecimento dos utilizadores que atuam em determinado domínio. Esse panorama se apresenta como desafio para os profissionais de Sistemas de Informação e de Ciência da Informação.

O perfil de aplicação de metadados proposto por este estudo procurou integrar os TNIs e a semântica de elementos de metadados no domínio da Oncologia no contexto dos artigos científicos em repositórios digitais científicos. Para tanto, requereu aportes teóricos e práticos das disciplinas de Sistemas de Informação e da Ciência da Informação. Este estudo demonstra a simbiose existente entre as duas áreas, em que as competências se mesclam, de forma a produzirem contributos para o desenvolvimento e a melhoria dos sistemas de informação.

Os profissionais que atuam nessas duas disciplinas podem compartilhar conhecimentos, ferramentas e métodos como o Design Science Research, a identificação de requisitos funcionais e informacionais, a

modelagem de domínio, a definição de modelo de dados e, sobretudo, os resultados, de modo a atender às necessidades dos utilizadores junto aos sistemas de informação.

Este estudo também explorou o domínio da Oncologia sob a perspectiva dos Sistemas de Informação e da Ciência da Informação, no que se refere aos tipos de necessidades de informação de profissionais que atuam como investigadores, professores e médicos.

9.2 DESENVOLVIMENTO DA INVESTIGAÇÃO

A proposta deste estudo é potencializar a interoperabilidade semântica por meio de uma aplicação que reflita níveis de granularidade de informação que venham contribuir para a qualidade da recuperação de artigos científicos em repositórios digitais. Para atingir esse objetivo, foram realizados estudos com a utilização de documentos, literatura científica e indivíduos, empregando técnicas e métodos distintos, como questionários, análise da literatura e documental e também ferramentas como NVIVO®, Vision®, CMAP®, SurveyMonkey® e o Me4MAP.

O embasamento teórico contribuiu para repensar o processo de comunicação científica na perspectiva de uma forma de acesso revolucionária – o acesso livre –, que disponibiliza o conhecimento sem barreiras comerciais ou de direitos autorais. Porém, mesmo com esse avanço, a descoberta de recursos nos repositórios institucionais ainda enfrenta sérios problemas, devido à reduzida interoperabilidade semântica. Nesse sentido, buscou-se estudar fundamentos da Web Semântica para compreender o novo paradigma dos *Linked Data* e as melhores formas de explorar as tecnologias semânticas.

O aprofundamento nessa temática revelou a possibilidade de avançar na perspectiva dos utilizadores, identificando os tipos de necessidades informacionais de médicos e investigadores ao buscar artigos científicos e convertendo esses tipos em elementos de metadados.

Coerente com o que tem sido preconizado pela comunidade de metadados, fez-se necessário o desenvolvimento de um perfil de aplicação de metadados – o Onco-MAP. O objetivo não foi definir um modelo engessado como base para o Onco-MAP, mas propor um perfil de aplicação para o domínio da Oncologia que contemple os TNIs e seja dinâmico, como tudo que se refere ao próprio domínio, que é muito susceptível às influências do tempo, dos avanços tecnológicos e das novas descobertas. Observa-se que a Oncologia ainda é um campo a ser bastante explorado, principalmente no que se refere à identificação dos TNIs dos seus profissionais.

Constatou-se, com o desenvolvimento deste estudo, que é possível ter um modelo genérico para artigos, em que são identificadas apenas propriedades referentes à descrição física dos recursos (título, autor, volume, número, editora) e um modelo para um domínio específico que contemple propriedades adicionais que aumentam a granularidade da descrição. Este último pode conter informações sobre métricas, notas, informações sobre conflito de interesses, nível de evidência, e especificidades dos tópicos abordados nas investigações – câncer, tratamento, diagnóstico, substâncias químicas, entre outros –, principalmente se forem armazenados em RDCs.

Vale salientar que a informação pode ser estruturada de forma que um domínio seja melhor representado, mesmo tendo ciência de que qualquer tentativa de modelagem de um domínio, ao estabelecer classes e propriedades, exige interpretações subjetivas que nem sempre serão consenso entre as pessoas (Stuart, 2016).

Os repositórios possuem características que permitem aumentar o nível de granularidade, como reflexo dos TNIs dos utilizadores, podendo acrescentar qualidade à descrição dos recursos ali armazenados (artigos, teses, dissertações, livros, entre outros). Os elementos de metadados empregados para a descrição desses recursos podem ser preenchidos de forma automática ou semiautomática, com a ajuda dos profissionais de informação. Por ser um sistema aberto, esses metadados ainda podem ser completados por meio de *crowdsourcing*.

O uso do Me4MAP viabilizou o desenvolvimento do Onco-MAP, na mesma perspectiva que uma bula pode orientar o paciente quanto a informações sobre composição, utilização, posologia e contraindicações de um medicamento e, assim, contribuir com o tratamento. A adequação do medicamento pode ser feita a partir do perfil do paciente e do tratamento. Assim, foi possível adequar o método às particularidades do Onco-MAP e adicionar a identificação dos requisitos informacionais, aspecto não contemplado pelo Me4MAP.

9.3 RESULTADO

Este estudo traz três principais contributos:

a) **a identificação de TNIs de investigadores e médicos para artigos científicos no domínio da Oncologia**. Estudos têm sido desenvolvidos com o objetivo de identificar os TNIs de médicos e investigadores (estes últimos de forma mais tímida) que atuam em algumas especialidades. A identificação dos TNIs, em sua maior parte, é derivada de questões que surgem a partir de

atendimentos clínicos. No que concerne à Oncologia, observou-se um número limitado de trabalhos sob as perspectivas dos investigadores e dos médicos. Este estudo identificou ainda a escassez de investigações envolvendo TNIs dos profissionais que atuam em Oncologia (médicos e investigadores) em relação às fontes de informação, nomeadamente os artigos científicos, o que resultou na proposta deste trabalho.

b) **O desenvolvimento de um PAM no domínio da Oncologia**. Alinhado à perspectiva da potencialidade da interoperabilidade semântica entre repositórios, o PAM é a opção que melhor traduz, na visão da autora a partir da literatura consultada, a possibilidade dos repositórios avançarem no paradigma *Linked Open Data*.

c) **Um esquema de metadados a partir dos TNIs de médicos e de investigadores no domínio da Oncologia**. Há poucos esquemas de metadados voltados para o domínio da Oncologia ou ainda pouco consolidados, de forma a atender o nível de granularidade que os utilizadores podem requerer.

São apresentadas também a identificação e a análise de PAMs no contexto dos RDCs e publicações científicas, além das características e pontos de acesso de bases de dados bibliográficas.

9.4 AS LIMITAÇÕES E OS TRABALHOS FUTUROS

No decorrer do estudo, houve algumas dificuldades com a ocorrência de variações terminológicas e conceituais dos temas investigados (perfil de aplicação de metadados e necessidade de informação, que podem ter contribuído para a omissão de estudos relevantes no período da pesquisa bibliográfica) e com o acesso aos investigados.

É notório o baixo número de respondentes dos *survey*s eletrônicos, independente do indivíduo investigado ou do domínio, muitas vezes justificado pela escassez de tempo para responder, pelo volume de questionários recebidos, pela complexidade do instrumento de coleta ou pela falta de conhecimento do assunto. Nesta tese, podem ser constatados três casos:

a) um índice de respostas foi obtido, apontado pela literatura como "bom" para o *survey*s (investigação com gestores de repositórios) – envio do questionário (SurveyMonkey) para 2.165 administradores de RDC e obtenção de 431 respostas;

b) o número de respondentes do *survey* aplicado aos médicos e investigadores foi baixo. Isso pode ser justificado, em parte, pelo tempo de preenchimento do questionário, pelas características dos respondentes em relação à disponibilidade e pela dificuldade de acesso aos endereços dos *e-mails*

dos entrevistados – aplicação de Survey com envio do questionário (SurveyMonkey) para 301 para médicos e investigadores que atuam em Portugal, obtenção de 59 respostas. Seleção de 39 respondentes do Norte de Portugal. Foram desconsiderados 20 questionários por não possuírem identificação da região e pelos respondentes atuarem em outras regiões do país. A utilização de *links* para acesso ao questionário restringiu a identificação de quantas pessoas realmente tiveram acesso ao questionário, e isso impossibilitou novo contato com os respondentes e o reenvio do questionário para complementar respostas, por exemplo; e

c) o estudo Delphi no processo de validação, com utilização de questionário eletrônico, também teve uma adesão dentro da média, de acordo com a literatura – envio de 22 *e-mails* para 22 bibliotecários(as), 8 concluíram o questionário. Esse resultado pode estar associado à complexidade do questionário.

Vale salientar que a participação nas investigações foi decorrente da boa vontade dos respondentes, já que não houve nenhum tipo de incentivo.

TRABALHOS FUTUROS

Pretende-se ampliar a amostra, com o objetivo de consolidar as informações obtidas e expandir o estudo para outros tipos de recursos, como teses e dissertações.

REFERÊNCIAS

Agoritsas, T., Merglen, A., Courvoisier, D. S., Combescure, C., Garin, N., Perrier, A., & Perneger, T. V. (2012). Sensitivity and predictive value of 15 PubMed search strategies to answer clinical questions rated against full systematic reviews. *Journal of Medical Internet Research*, 14(3). Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3414859/.

Aitken, C., Power, R., & Dwyer, R. (2008). A very low response rate in an on-line *survey* of medical practitioners. *Australian and New Zealand Journal of Public Health*, 32(3), 288–289.

Aken, J. E. van. (2004). Management research based on the paradigm of the design sciences: the quest for field-tested and grounded technological rules. *Journal of Management Studies*, 41(2), 219–246.

Allen, M., Currie, L. M., Graham, M., Bakken, S., Patel, V. L., & Cimino, J. J. (2003). The classification of clinicians' information needs while using a clinical information system. In *AMIA Annual Symposium Proceedings* (p. 26). American Medical Informatics Association. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1480189/.

Alvarenga, L. (2003). Representação do conhecimento na perspectiva do ciência da informação em templ e espaço digitaisRepresentation of the knowledge in the perspective of the science of the information in digital time and space. *Encontros Bibli: Revista Eletrônica de Biblioteconomia E Ciência Da Informação*, 8(15).

Amato, G., & Straccia, U. (1999). User profile modeling and applications to digital libraries. In *Research and Advanced Technology for Digital Libraries* (pp. 184–197). Springer. Retrieved from http://*link*.springer.com/chapter/10.1007/3-540-48155-9_13.

Ambinder, D. M., & Marcondes, C. H. (2012). Artigos científicos digitais na Web: novas experiências para apresentação, acesso e leitura. In *Anais do XIII Encontro Nacional de Pesquisa em Ciência da Informação* (pp. 1–20). Rio de Janeiro: ANCIB.

American Society of Clinical Oncology. (2016). State of Cancer Care in America: 2015 Hill Briefing Webcast. American Society of Clinical Oncology. Retrieved from http://www.asco.org/practice-research/state-cancer-care-america-2015-hill-briefing-webcast.

Andrade, M., & Baptista, A. A. (2014). Researchers' information needs in the bibliographic database: A literature review. *Information Services & Use*, 34(Special Issue), 241–248. https://doi.org/10.3233/ISU-140744.

Andrade, M. C. de, & Baptista, A. A. (2015). The use of application profiles and metadata schemas by digital repositories: findings from a *survey*. Retrieved from http://dspace2.ufes.br/handle/10/2022.

Antezana, E., Kuiper, M., & Mironov, V. (2009). Biological knowledge management: the emerging role of the Semantic Web technologies. *Briefings in Bioinformatics*, 10(4), 392–407. https://doi.org/10.1093/bib/bbp024.

Armbruster, C., & Romary, L. (2010). Comparing repository types. Retrieved from http://arxiv.org/ftp/arxiv/papers/1005/1005.0839.pdf.

Associação Brasileira de Normas Técnicas. (2012a). *NBR 6024: informação e documentação - numeração progressiva das seções de um documento - apresentação*. Rio de Janeiro: ABNT.

Associação Brasileira de Normas Técnicas. (2012b). *NBR 6027: informação e docuemntação - sumário - apresentação*. Rio de Janeiro: ABNT.

Associação Brasileira de Normas Técnicas. (2012c). *NBR 14724: informação e documentação - trabalhos acadêmicos - apresentação*. Rio de Janeiro: ABNT.

Attribute. In: Glossary of unified modeling language terms. *Wikipedia The Free Encyclopedia*. Available in https://en.wikipedia.org/wiki/Glossary_of_Unified_Modeling_Language_terms

Attwood, T., Kell, D., McDermott, P., Marsh, J., Pettifer, S., & Thorne, D. (2009). Calling international rescue: knowledge lost in literature and data landslide! *Biochem. J*, 424, 317–333.

Avison, D., & Myers, M. (2005). Qualitative research. In D. Avison & J. Pries-Heje (Eds.), *Research in information systems:* a handbook for research supervisors and their students (pp. 239–253). Amsterdam: Elsevier.

Baker, T., Bermès, E., Coyle, K., Dunsire, G., Isaac, A., Murray-Rust, P., ... Singer, R. (2011). Library *Linked Data* Incubator Group Final Report. Retrieved May 8, 2013, from http://www.w3.org/2005/Incubator/lld/XGR-lld-20111025/.

Baptista, A. A. (2002). *Informattica online*: um enquadramento para a publicação em linha de revistas científicas electrónicas. Retrieved from https://repositorium.sdum.uminho.pt/handle/1822/284.

Baptista, A. A. (2010). A falar nos entendemos: a interoperabilidade entre repositórios digitais. In: Gomes, M. J., Rosa, F. (orgs.). *Repositórios institucionais*: democratizando o acesso ao conhecimento. Salvador: EDUFBA. Retrieved from http://repositorium.sdum.uminho.pt/handle/1822/11517.

Baptista, A. A., Costa, S. M. de S., Kuramoto, H., & Rodrigues, E. (2007). Comunicação científica: o papel da Open Archives Initiative no contexto do acesso livre. Retrieved from http://repositorio.unb.br/handle/10482/635.

Baptista, S. G., & Cunha, M. B. da. (2007). Estudo de usuários: visão global dos métodos de coleta de dados. *Perspectivas Em Ciência Da Informação*, 12(2), 168–184.

Baptista, A. A., & Ferreira, M. (2007). Conversas de repositório. *BJIS*, Marília, v. 0. Retrieved from http://www.doaj.org/doaj?func=fulltext&aId=1423654.

Bates, M. J. (1996). The Getty end-user online searching project in the humanities: report no. 6: overview and conclusions. *College & Research Libraries*, 57(6), 514–23.

Bax, M. P. (2014). Design science: filosofando da pesquisa em Ciência da Informação e Tecnologia. In *XV ENANCIB "Além das nuvens*. Belo Horizonte: ANCIB. Retrieved from http://www.bax.com.br/publications/artigos/design-science-filosofia-da-pesquisa-em-ciencia-da-informacao-e-tecnologia/view

Bechhofer, S., van Harmelen, F., Hendler, J., Horrocks, I., McGuinness, D. L., Patel-Schneider, P. F., & Stein, L. A. (2004, February 10). OWL web ontology language reference. Retrieved April 28, 2013, from http://www.w3.org/TR/owl-ref/.

Bekaert, J., & Van De Sompel, H. (2006). Augmenting interoperability across scholarly repositories. Retrieved from https://www.mysciencework.com/publication/show/76125d0553108a0e587585b75b09e6eb.

Bell, D. (2016, dezembro). Fundamentos básicos de UML: o diagrama de classes: uma introdução aos diagramas de estrutura em UML 2. IBM. Retrieved from

https://www.ibm.com/developerworks/br/rational/library/content/RationalEdge/sep04/bell/index.html

Berghel, H., Berleant, D., Foy, T., & McGuire, M. (1999). Cyberbrowsing: information customization on the Web. *Journal of the American Society for Information Science*, 50(6), 505–513. https://doi.org/10.1002/(SICI)1097-4571(1999)50:6<505::AID-ASI5>3.0.CO;2-R.

Berners-Lee, T. (2009, June 18). *Linked Data* - design issues. Retrieved April 9, 2013, from http://www.w3.org/DesignIssues/*Linked Data*.html.

Berners-Lee, T., Hendler, J., & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), 28–37.

Berners-Lee, T., Cailliau, R., Groff, J.F., & Pollermann, B. (1992). World-wide web: the information universe. *Internet Research*, *20*(4), 461–471. https://doi.org/10.1108/10662241011059471.

Bibliographic database - New World Encyclopedia. (n.d.). Retrieved May 25, 2016, from http://www.newworldencyclopedia.org/entry/Bibliographic_database.

Bishop, A. P. (1998). Digital libraries and knowledge disaggregation: the use of journal article components. In *Proceedings of the third ACM conference on Digital libraries* (pp. 29–39). Retrieved from http://dl.acm.org/citation.cfm?id=276679.

Bishop, A. P. (1999). Document structure and digital libraries: how researchers mobilize information in journal articles. *Information Processing & Management*, 35(3), 255–279.

Bizer, C., Heath, T., & Berners-Lee, T. (2009). *Linked Data*: the story so far. *International Journal on Semantic Web and Information Systems*, 5(3), 1–22.

Bjerre, L. M., Paterson, N. R., McGowan, J., Hogg, W., Campbell, C. M., Viner, G., & Archibald, D. (2013). What do primary care practitioners want to know? A content analysis of questions asked at the point of care. *Journal of Continuing Education in the Health Professions*, 33(4), 224–234.

Björk, B.-C. (2007). A model of scientific communication as a global distributed information system. *Information Research*, 12(2), 1–48. Recuperado em: http://docs.lib.purdue.edu/cgi/viewcontent.cgi?article=1797&context=iatul.

Blaha, M., & Rumbaugh, J. (2006). *Modelagem e projetos baseados em objetos com UML* (2. ed.). Rio de Janeiro: Elsevier.

Borgman, C. L. (1986). Why are online catalogs hard to use? Lessons learned from information-retrieval studies. *Journal of the American Society for Information Science*, 37(6), 387–400.

Borgman, C. L. (2007). The continuity of scholarly communication. In *Scholarship in the Digital Age: Information*, *Infrastructure*, *and the Internet* (pp. 46–76). Cambridge, MA: MIT Press.

Borko, H. (1968). Information science: what is it? *American Documentation*, 19(1), 3–5.

Bowden, V. M., Kromer, M. E., & Tobia, R. C. (1994). Assessment of physicians' information needs in five Texas counties. *Bulletin of the Medical Library Association*, 82(2), 189.

Brasil. Ministério da Saúde. (2013). *Glossário temático:* controle de câncer*.* Brasília.

Brasil. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos, Departamento de Ciência e Tecnologia. (2015). Síntese de Evidências para Políticas de Saúde: estimulando o uso de evidências científicas na tomada de decisão. Ministério da Saúde.

Bryant, S. L. (2000). The information needs and information seeking behaviour of family doctors: a selective literature review. *Health Libraries Review*, 17(2), 83–90.

Bryant, S. L. (2004). The information needs and information seeking behaviour of family doctors. *Health Information & Libraries Journal*, 21(2), 84–93.

Calva-González, J. J. (2003). Las necesidades de información de los investigadores del área de Humanidades y Ciencias Sociales. *Revista General de Información Y Documentación*, 13(2), 155–180.

Campbell, D. M., Redman, S., Jorm, L., Cooke, M., Zwi, A. B., & Rychetnik, L. (2009). Increasing the use of evidence in health policy: practice and views of policy makers and researchers. *Australia and New Zealand Health Policy*, 6(1), 21.

Campos, M. de A., & Gomes, H. E. (2008). Taxonomia e classificação: o princípio de categorização. *DataGramaZero: Revista de Ciência da Informação*, 9(4).

Caplan, P. (2009). Understanding Premis. *D-Lib Magazine*, 15(3/4). Retrieved from http://www.citeulike.org/group/14155/article/8289782.

Castells, P. (2003). La web semántica. *Sistemas Interactivos Y Colaborativos En La Web*, 195–212.

Catarino, M. E. (2009). *Integração das folksonomias nos metadados:* identificação de novos elementos como contributo para a descrição de recursos em repositórios (Tese). Universidade do Minho, Guimarães, Portugal.

Catarino, M. E., & Baptista, A. A. (2008). Social tagging and Dublin Core: a preliminary proposal for an application profile fo DC social tagging. Retrieved from http://repositorium.sdum.uminho.pt/handle/1822/8361.

Centro de Combate ao Câncer. (2014). O que é oncologia. Retrieved from http://www.cccancer.net/site/index.php/o-que-e-oncologia-2/.

Chalmers, I. (1995). What do I want from health research and researchers when I am a patient? *BMJ: British Medical Journal*, 310(6990), 1315.

Chan, L. M., & Zeng, M. L. (2006). Metadata interoperability and standardization-A study of methodology, Part II. *D-Lib Magazine*, 12(6), 4.

Cheng, G. Y. (2004). A study of clinical questions posed by hospital clinicians. *Journal of the Medical Library Association*, 92(4), 445.

Choo, C. W. (2003). A organização do conhecimento: como as organizações usam a informação para criar significado, construir conhecimento e tomar decisões. Senac São Paulo. Retrieved from https://pt-static.z-dn.net/files/d2a/6b22b57802f85a2c74fd96a4e712d11b.doc.

Cimino, J. J. (1996). Linking patient information systems to bibliographic resources. *Methods of Information in Medicine-Methodik Der Information in Der Medizin*, 35(2), 122–126.

Cochrane, A. L. (1972). Effectiveness and efficiency: Random reflections on health services. *London: Nuffield Provincial Hospitals Trust*.

Cogdill, K. W., Friedman, C. P., Jenkins, C. G., Mays, B. E., & Sharp, M. C. (2000). Information needs and information seeking in community medical education. *Academic Medicine*, *75*(5), 484–486.

Collins, S. A., Currie, L. M., Bakken, S., & Cimino, J. J. (2009). Information needs, Infobutton Manager use, and satisfaction by clinician type: a case study. *Journal of the American Medical Informatics Association*, 16(1), 140–142.

Costa, R. M. P. (2010). Luta contra o cancro e oncologia em Portugal: estruturação e normalização de uma área científica (1839-1974). Retrieved from https://repositorio-aberto.up.pt/handle/10216/50414.

Costa, S. M. de S. (2008). Abordagens, estratégias e ferramentas para o acesso aberto via periódicos e repositórios institucionais em instituições acadêmicas brasileiras. Retrieved from http://repositorio.unb.br/handle/10482/16048.

Costa, M. M. (2012). As bibliotecas brasileiras em 2018: resultados da técnica delfos. *Perspectivas em Ciência da Informação*, 17(1): 74-93.

Council, N. R. (2011). *Toward precision medicine: building a knowledge network for biomedical research and a new taxonomy of disease*. Washington: National Academies Press. Retrieved from https://books.google.com.br/books?hl=pt-BR&lr=&id=vcJABAAAQBAJ&oi=fnd&pg=PR1&dq=Toward +Precision+Medicine:+Building+a+Knowledge+Network+for+Biomedical+Research+and+a+New+Taxonomy+of+Disease&ots=-nwHwWEx80&sig=54o7KeuUF72OgEmBWTEemb5sepg.

Courtright, C. (2007). Context in information behavior research. *Annual Review of Information Science and Technology*, 41(1), 273–306.

Coyle, K., & Baker, T. (2009). Guidelines for Dublin Core application profiles. Retrieved from http://dublincore.org/documents/2009/05/18/profile-guidelines/

Crawford, S. (1978). Information needs and uses. *Annual Review of Information Science and Technology*, 13(1), 61.

Creswell, J. W. (2009). Qualitative methods. In *Research design:* qualitative, quantitative, and mixed methods approaches (3rd ed., p. cap. 9.). Thousand Oaks: Sage Publications.

Creswell, J. W. (2012). *Qualitative inquiry and research design:* choosing among five approaches. Sage publications. Retrieved from https://books.google.com.br/books?hl=pt-BR&lr=&id=OJYEbDtkxq8C&oi =fnd&pg=PP1&dq=Qualitative+inquiry+and+research +design:&ots=eO9d8OOngq&sig=h0oKwdfXZjVwPeZlXnCp7jrG2CE.

Crow, R. (2002). The case for institutional repositories: a SPARC position paper. *ARL Bimonthly Report 223*. Retrieved from http://works.bepress.com/ir_research/7.

Crowston, K., & Kwasnik, B. H. (2003). Can document-genre metadata improve information access to large digital collections. *Library Trends*, 52(2), 345–361.

Cull, W. L., O'Connor, K. G., Sharp, S., & Tang, S. S. (2005). Response rates and response bias for 50 *surveys* of pediatricians. *Health Services Research*, 40(1), 213–226. https://doi.org/10.1111/j.1475-6773.2005.00350.x.

Dalmoro, M., Vieira, K. M. (2013). Dilemas na construção de escalas tipo Likert: o número de itens e a disposição influenciam nos resultados? *RGO Revista Gestão Organizacional.* 6 (Esp.):161-74.

Darby, R. M., Jones, C. M., Gilbert, L. D., & Lambert, S. C. (2008). Increasing the productivity of interactions between subject and institutional repositories. *New Review of Information Networking*, 14(2), 117–135. https://doi.org/10.1080/13614570903359381.

Davies, K. (2008). *Clinical information needs of doctors in the UK*. Retrieved from https://dspace.lboro.ac.uk/dspace-jspui/handle/2134/8089.

Day, J., & Bobeva, M. (2005). A generic toolkit for the successful management of Delphi studies. *The Electronic Journal of Business Research Methodology*, 3(2), 103–116.

De Azambuja, E., Ameye, L., Paesmans, M., Zielinski, C. C., Piccart-Gebhart, M., & Preusser, M. (2014). The landscape of medical oncology in Europe by 2020. *Annals of Oncology*, 25(2), 525–528.

De Villiers, M. R. (2012). Models for Interpretive Information Systems Research, Part 2: Design Research, Development. *Research Methodologies*, *Innovations and Philosophies in Software Systems Engineering and Information Systems*, 238.

Dee, C., & Blazek, R. (1993). Information needs of the rural physician: a descriptive study. *Bulletin of the Medical Library Association*, 81(3), 259.

Del Fiol, G., Workman, T. E., & Gorman, P. N. (2014). Clinical questions raised by clinicians at the point of care: a systematic review. *JAMA Internal Medicine*, 174(5), 710–718.Delbecq, A. L., Van de Ven, A. H., & Gustafson, D. H. (1986). *Group techniques for program planning :* a guide to nominal group and delphi processes. Middleton: Green Briar.

Demografia Médica no Brasil 2015. (2015). São Paulo: Universidade de São Paulo. Retrieved from https://www.google.pt/?gws_rd=ssl#q=Demografia+M%C3%A9dica+no+Brasil+2015

Devey, M., & Côté, M.-C. (2006). The development and use of metadata application profiles. *The Serials Librarian*, 51(2), 103–115. https://doi.org/10.1300/J123v51n02_08

Di Iorio, A. (s.d.). Um modelo de dados para metadados de preservação. Retrieved May 6, 2013, from www.digitalpreservationeurope.eu/publications/.../PORTODIIORIO.pdf.

Diamantopoulos, N., Sgouropoulou, C., Kastrantas, K., & Manouselis, N. (2011). Developing a metadata application profile for sharing agricultural scientific and scholarly research resources. *Metadata and Semantic Research*, 453–466.

Digital Repository Infrastructure Vision For European Research. (2009). Directrizes DRIVER 2.0: directrizes para fornecedores de conteúdos: exposição de recursos textuais com o protocolo OAI-PMH. Retrivied http//:www.driver-support.eu/documents/DRIVER_Guidelines_v2_ Final__ PT.pdf.

Dinakarpandian, D., Lee, Y., Vishwanath, K., & Lingambhotla, R. (2006). MachineProse: an ontological framework for scientific assertions. *Journal of the American Medical Informatics Association*, 13(2), 220–232.

Direcção-Geral da Saúde. Direcção de Serviços de Planeamento. (2012). *Rede de Referenciação Hospitalar de Oncologia.* Lisboa: Direcção-Geral de Saúde. Retrieved from https://www.google.pt/search?q=Rede+de+Referencia%C3%A7%C3%A3o+HospitalardeOncologia&ie=utf-8&oe=utf-8&gws_rd=cr&ei=pOebVZWPN8zWUdDQragL

Divoli, A., Wooldridge, M. A., & Hearst, M. A. (2010). Full text and figure display improves bioscience literature search. *PloS One*, 5(4), e9619.

Dogan, R. I., Murray, G. C., Névéol, A., & Lu, Z. (2009). Understanding PubMed® user search behavior through log analysis. *Database: The Journal of Biological Databases and Curation*, *2009*. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2797455/

Drake, M. (2004). Institutional repositories hidden treasures. *Searcher-Medford Nj-*, 12(5), 41–45.

Drolet, B. C., & Lorenzi, N. M. (2011). Translational research: understanding the continuum from bench to bedside. *Translational Research*, 157(1), 1–5.

Dublin Core Metadata Iniciative. (2012a). DCMI grammatical principles. Retrieved from http://dublincore.org/usage/documents/principles/.

Dublin Core Metadata Iniciative. (2012b). DCMI metadata terms. Retrieved from http://dublincore.org/documents/dcmi-terms.

Duerst, M., & Suignard, M. (2005). Internationalized Resource Identifiers (IRIs). Retrieved October 19, 2016, from https://tools.ietf.org/html/rfc3987.

Duval, E., Hodgins, W., Sutton, S., & Weibel, S. L. (2002). Metadata principles and practicalities. *D-Lib Magazine*, 8(4), 16.

Ebell, M. H., Cervero, R., & Joaquin, E. (2011). Questions asked by physicians as the basis for continuing education needs assessment. *Journal of Continuing Education in the Health Professions*, 31(1), 3–14.

Enoksson, F. A MoinMoin Wiki Syntax for Description Set Profiles (s.d.). Retrieved http://dublincore.org/ documents/dsp-wiki-syntax/.

Ely, J. W., Burch, R. J., & Vinson, D. C. (1992). The information needs of family physicians: case-specific clinical questions. *The Journal of Family Practice*. Retrieved from http://psycnet.apa.org/psycinfo/1993-10818-001.

Ely, J. W., Osheroff, J. A., Ebell, M. H., Bergus, G. R., Levy, B. T., Chambliss, M. L., & Evans, E. R. (2000b). Analysis of questions asked by family physicians regarding patient care. *Western Journal of Medicine*, *172*(5), 315.

Ely, J. W., Osheroff, J. A., Gorman, P. N., Ebell, M. H., Chambliss, M. L., Pifer, E. A., & Stavri, P. Z. (2000a). A taxonomy of generic clinical questions: classification study. *BMJ*, 321(7258), 429–432. https://doi.org/10.1136/bmj.321.7258.429.

European Comission. (2008). Interoperable Delivery of European eGovernment Services to public Administrations, Businesses and Citizens – IDABC. In *Draft for public comments: as basis for EIF 2.0*. Bruxelles: European Comission. Retrieved from : http://ec.europa.eu/idabc/servlets/Docb0db.pdf?id=31597

European Comission. (2010). Interoperability solutions for European Public Administrations. European Comission. Bruxelles: European Comission. Retrieved from: ec.europa.eu/isa/documents/isa_annex_ii_eif_en.pdf.

Falagas ME, Pitsouni EI, Malietzis GA, Pappas G. (2008) Comparison of PubMed, Scopus,Web of Science, and Google Scholar: strengths and weaknesses. *FASEB Journal*, Feb;22(2):338-42

Falconer, D. J., & Hodgett, R. A. (1999). Why executives don't respond to your *survey*. In *Proceedings of the 10th Australasian Conference on information Systems* (pp. 279–285). Citeseer. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.2.5712 &rep=rep1&type=pdf.

Fan, W., & Yan, Z. (2010). Factors affecting response rates of the web *survey*: a systematic review. *Computers in Human Behavior*, 26(2), 132–139.

Fidel, R. (1993). Qualitative methods in information retrieval research. *LISR*, 15, 219–247.

Fink, A., Kosecoff, J., Chassin, M., & Brook, R. H. (1984). Consensus methdos: characteristics and guidelines for use. *American Journal of Public Health*, *74*(9), 979–983.

Figueiredo, N. M. de. (1983). Aspectos especiais de estudos de usuários. *Ciência da Informação*, 12(2). Retrieved from http://revista.ibict.br/ciinf/index.php/ciinf/article/viewArticle/1496.

Fischbacher, C., Chappel, D., Edwards, R., & Summerton, N. (2000). Health *survey*s via the Internet: quick and dirty or rapid and robust? *Journal of the Royal Society of Medicine*, 93(7), 356.

Forsythe, D. E. (1998). Using ethnography to investigate life scientists' information needs. *Bulletin of the Medical Library Association*, 86(3), 402.

Forsythe, D. E., Buchanan, B. G., Osheroff, J. A., & Miller, R. A. (1992). Expanding the concept of medical information: an observational study of physicians' information needs. *Computers and Biomedical Research*, 25(2), 181–200.

Foulonneau, M., Cole, T. W., Habing, T. G., & Shreeves, S. L. (2005). Using collection descriptions to enhance an aggregation of harvested item-level metadata (p. 32). ACM Press. https://doi.org/10.1145/1065385.1065393.

Fourie, I., & Bakker, S. (2009). Exploring information needs and information behaviour in oncology contexts: the importance of tasks, roles and organisational settings. In *Positioning the Profession: the Tenth International Congress on Medical Librarianship* (pp. 1–15). Retrieved from http://espace.library.uq.edu.au/view/UQ:179911.

Frauenfelder, M. (2001). A smarter Web. Retrieved April 9, 2017, from https://www.technologyreview.com/s/401241/a-smarter-web/

Freeman Rohde, N. (1986). Information needs. *Advances in Librarianship*, 14, 49–73.

Frei, E. (1982). Clinical cancer research: an embattled species. *Cancer*, 50(10), 1979–1992.

Furlan, J. D. (1998). *Modelagem de objetos através da UML*. São Paulo: Makron Books.

Fusco, E. (2010). *Modelos conceituais de dados como parte do processo da catalogação*: perspectiva de uso dos FRBR no desenvolvimento de catálogos bibliográficos digitais. (Tese). Universidade Estadual Paulista, Marília. Retrieved from http://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/fusco_e_do_mar.pdf

Galvão, M. C. B., & Ricarte. (2013, June 27). A SNOMED CT e os sistemas de informação em saúde. Retrieved May 30, 2017, from http://www.ofaj.com.br/colunas_conteudo.php?cod=757.

Garvey, W. D., & Griffith, B. C. (1979). Scientific communication as a social system. *Communication: The Essence of Science*, 148–164.

Gibbons, P., Arzt, N., Burke-Beebe, S., Chute, C., Dickinson, G., Flewelling, T., ... Stanford, J. (2007). Coming to terms: Scoping interoperability for health care. Retrieved from http://www.citeulike.org/group/15536/article/10705562.

Giuse, N. B., Huber, J. T., Giuse, D. A., Brown, C. W., Bankowitz, R. A., & Hunt, S. (1994). Information needs of health care professionals in an AIDS outpatient clinic as determined by chart review. *Journal of the American Medical Informatics Association*, 1(5), 395–403.

González-González, A. I., Dawes, M., Sánchez-Mateos, J., Riesgo-Fuertes, R., Escortell-Mayor, E., Sanz-Cuesta, T., & Hernández-Fernández, T. (2007). Information needs and information-seeking behavior of primary care physicians. *The Annals of Family Medicine*, 5(4), 345–352.

González-González, A. I., Escortell Mayor, E., Hernández Fernández, T., Sánchez Mateos, J. F., Sanz Cuesta, T., & Riesgo Fuertes, R. (2005). Necesidades de información de los médicos de atención primaria: análisis de preguntas y su resolución. *Atención Primaria*, 35(8), 419–422.

Gorman, P. (1993). Does the medical literature contain the evidence to answer the questions of primary care physicians? Preliminary findings of a study. In *Proceedings of the Annual Symposium on Computer Application in Medical Care* (p. 571). American Medical Informatics Association. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2850641/.

Gorman, P. N. (1995). Information needs of physicians. *Journal of the American Society for Information Science*, *46*(10), 729–736.

Gorman, P. N., Ash, J., & Wykoff, L. (1994). Can primary care physicians' questions be answered using the medical journal literature? *Bulletin of the Medical Library Association*, 82(2), 140.

Graber, M. A., Randles, B. D., Monahan, J., Ely, J. W., Jennissen, C., Peters, B., & Anderson, D. (2007). What questions about patient care do physicians have during and after patient contact in the ED? The taxonomy of gaps in physician knowledge. *Emergency Medicine Journal*, 24(10), 703–706.

Grau, B. C., Horrocks, I., Motik, B., Parsia, B., Patel-Schneider, P., & Sattler, U. (2008). OWL 2: The next step for OWL. *Web Semantics:* Science, Services and Agents on the World Wide Web, 6(4), 309–322. https://doi.org/10.1016/j.websem.2008.05.001.

Grefsheim, S. F., & Rankin, J. A. (2007). Information needs and information seeking in a biomedical research setting: a study of scientists and science administrators. *Journal of the Medical Library Association:* JMLA, 95(4), 426.

Gregor, S., & Hevner, A. R. (2013). Positioning and presenting design science research for maximum impact. *MIS Quarterly*, 37(2), 337–356.

Grimm, P. (2010). Pretesting a Questionnaire. *Wiley International Encyclopedia of Marketing*. Retrieved from http://onlinelibrary-wiley-com.ez43.periodicos.capes.gov.br/doi/10.1002/9781444316568.wiem02051/full

Groth, P., Gibson, A., & Velterop, J. (2010). The anatomy of a nanopublication. *Information Services and Use*, 30(1), 51–56.

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, 5(2), 199–220.

Guimarães, R. (2013). Pesquisa Translacional: uma interpretação. *Cienc Saude Coletiva*, 18(6), 1731–44.

Hajdu, S. I. (2011). A note from history: Landmarks in history of cancer, part 1. *Cancer*, 117(5), 1097–1102. https://doi.org/10.1002/cncr.25553.

Hanief Bhat, M. (2010). Interoperability of open access repositories in computer science and IT-an evaluation. *Library Hi Tech*, *28*(1), 107–118.

Harnad, S. (2007). The green road to open access: a leveraged transition. Retrieved from http://eprints.ecs.soton.ac.uk/15753.

Harnad, S., Brody, T., Vallières, F., Carr, L., Hitchcock, S., Gingras, Y., ... Hilf, E. R. (2008). The access/impact problem and the Green and Gold Roads to Open Access: An Update. *Serials Review*, 34(1), 36–40. https://doi.org/10.1016/j.serrev.2007.12.005

Haynes, R. B., Wilczynski, N., McKibbon, K. A., Walker, C. J., & Sinclair, J. C. (1994). Developing optimal search strategies for detecting clinically sound studies in MEDLINE. *Journal of the American Medical Informatics Association*, 1(6), 447–458.

He, K.-Q., Wang, J., & Liang, P. (2010). Semantic interoperability aggregation in service requirements refinement. *Journal of Computer Science and Technology*, 25(6), 1103–1117.

Heery, R., & Anderson, S. (2005). Digital repositories review. Retrieved from http://opus.bath.ac.uk/23566/2/digital-repositories-review-2005.pdf.

Heery, R. (2009). Digital repositories roadmap review: towards a vision for research and learning in 2013. Retrieved from http://kennison.name/files/zopestore/uploads/libraries/documents/reproadmapreviewfinal.pdf.

Heery, R., & Patel, M. (2000). Application profiles: mixing and matching metadata schemas. *Ariadne*, 25, 27–31.

Hevner, A. R. (2007). A three cycle view of design science research. *Scandinavian Journal of Information Systems*, 19(2), 4.

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research. Retrieved from http://ebooks.cawok.pro/John.Wiley.and.Sons.Information.Systems.The.State.of.the.Field.May. 2006.eBook-LinG.pdf#page=218.

Hillmann, D. Using Dublin Core. Metadata Dublin Core® Metadata Initiative. Retrieved http://dublincore.org/documents/usageguide/.

História da Medicina. (n.d.). Retrieved July 7, 2015, from http://www.suapesquisa.com/ecologiasaude /medicina.htm.

Hjørland, B., Nielsen, L. K., & Williams, M. E. (2001). Subject access points in electronic retrieval. *Annual Review of Information Science and Technology*, *35*, 249–298.

Hodge, G. M. (2000). An information life-cycle approach: best practices for digital archiving. *Journal of Electronic Publishing*, 5(4). https://doi.org/http://dx.doi.org/10.3998/3336451.0005.406.

Hoffman, G. L. (2009). Meeting users' needs in cataloging: what is the right thing to do? *Cataloging & Classification Quarterly*, 47(7), 631–641.

Hurd, J. (2000). The transformation of scientific communication: A model for 2020. *Journal Of The American Society For Information Science*, 51(14), 1279–1283.

IBICT (n.d.) Sobre Repositórios Digitais — IBICT. Retrieved April 14, 2017, from http://www.ibict.br/informacao-para-ciencia-tecnologia-e-inovacao%20/repositorios-digitais.

International Digital Enterprise Alliance. (2013). About PRISM. Retrieved May 6, 2013, from http://www.prismstandard.org/about/.

Jansen, B. J. (2009). The methodology of search log analysis. In *Handbook of research on web log analysis*. New York: Information Science Reference.

Jenal, S., Vituri, D. W., Ezaías, G. M., Silva, L. A. da, Caliri, M. H. L. (2012). O processo de revisão por pares: uma revisão integrativa de literatura. *Acta Paulista de Enfermagem*, *25*(5), 802–808.

Jerome, R. N., Giuse, N. B., Gish, K. W., Sathe, N. A., & Dietrich, M. S. (2001). Information needs of clinical teams: analysis of questions received by the Clinical Informatics Consult Service. *Bulletin of the Medical Library Association*, 89(2), 177.

Johnson, P. T., Chen, J. K., Eng, J., Makary, M. A., & Fishman, E. K. (2008). A comparison of world wide web resources for identifying medical information. *Academic Radiology*, 15(9), 1165–1172.

Julien, H., & Duggan, L. J. (2000). A longitudinal analysis of the information needs and uses literature. *Library & Information Science Research*, 22(3), 291–309.

Kayo, E. K., & Securato, J. R. (1997). Método Delphi: fundamentos, críticas e vieses. *Cad Pesq Admin*, 1(4), 51–61.

Keeney, S., Hasson, F., & McKenna, H. (2001). Research guidelines for the Delphi *survey* technique. *Journal of Advanced Nursing*, *32*(4), 1008–1015.

Kerlinger, FN. (1980). *Metodologia da pesquisa em ciências sociais:* um tratamento conceitual. Sao Paulo: EDUSP.

Kellerman, S. E., & Herold, J. (2001). Physician response to *survey*s. *American Journal of Preventive Medicine*, 20(1), 61–67. https://doi.org/10.1016/S0749-3797(00)00258-0.

Khoo, M., Rozaklis, L., & Hall, C. (2012). A *survey* of the use of ethnographic methods in the study of libraries and library users. *Library & Information Science Research*, 34, 82–91. https://doi.org/10.1016/j.lisr.2011.07.010.

Kim, S. (2010a). A conceptual framework of information requirements for scientists using human biological samples. *Information Research*, 15(1), 1–26.

Kim, S. (2010b). The information needs of biomedical scientists using biological samples: a conceptual framework. *Information Research*, 15(1). Retrieved from http://nile.lub.lu.se/ojs/index.php/infores/article/view/1655

Kim, S., & Gilbertson, J. (2007). Information requirements of cancer center researchers focusing on human biological samples and associated data. *Information Processing & Management*, 43(5), 1383–1401.

Knublauch, H., & Kontokostas, D. (2017, February 2). Shapes Constraint Language (SHACL). Retrieved February 9, 2017, from https://www.w3.org/TR/shacl/.

Knublauch, H., TopQuadrant, Inc., Prud'hommeaux, E., & W3C/MIT. (2015, April 2). Shapes Constraint Language (SHACL). W3C Editor. Retrieved from https://w3c.github.io/data-shapes/shacl/.

Kôche, J. C. (1996). Fundamentos de metodologia científica, Caxias do Sul/Porto Alegre, EDUCS/EST. *Vozes*, (14).

Komatsoulis, G. A., Warzel, D. B., Hartel, F. W., Shanbhag, K., Chilukuri, R., Fragoso, G., ... others. (2008). caCORE version 3: Implementation of a model driven, service-oriented architecture for semantic interoperability. *Journal of Biomedical Informatics*, *41*(1), 106–123.

Koopmans, N. I. (2002). What's your question? The need for research information from the perspective of different user groups. Retrieved from http://dspacecris.eurocris.org/jspui/handle/123456789/148.

Korjonen-Close, H. (2005). The information needs and behaviour of clinical researchers: a user-needs analysis. *Health Information And Libraries Journal*, *22*(2).

Kuruppu, P. U., & Gruber, A. M. (2006). Understanding the Information Needs of Academic Scholars in Agricultural and Biological Sciences. *The Journal of Academic Librarianship*, 32(6), 609–623. https://doi.org/10.1016/j.acalib.2006.08.001.

Lagoze, C., Van de Sompel, H., Nelson, M., & Warner, S. (2002, June 14). Open Archives Initiative - Protocol for Metadata Harvesting - v.2.0. Retrieved May 8, 2013, from http://www.openarchives.org/OAI/openarchivesprotocol.html.

Lakatos, E. M., & Marconi, M. A. (2003). *Fundamentos de metodologia científica* (5th ed.). São Paulo: Atlas.

Lang, M. (2002). The use of web-based international *survey*s in information systems research. Retrieved from https://aran.library.nuigalway.ie/handle/10379/266.

Lassila, O., & Swick, R. R. Resource description framework (RDF) model and syntax. Retrieved from http://www.w3.org/TR/WD-rdf-syntax-971002.

Le Codiac, Y.-F. (1996). *A ciência da informação*. Brasília, DF: Briquet de Lemos/Livros.

Lee, J. H., & Downie, J. S. (2004). Survey of music information needs, uses, and seeking behaviours: preliminary findings. In *ISMIR* (Vol. 2004, p. 5th). Retrieved from http://people.lis.illinois.edu/~jdownie/ismir2004_*survey*_downie_draft.pdf.

Lee, P. J. (2013). The information needs of clinicians: a study of the doctors Nova Scotia Clinical Library. *Journal of the Canadian Health Libraries Association*, *34*(3), 159–163. https://doi.org/10.5596/c13-058.

Leite, J. C. S. P. (1994). *Engenharia de requisitos*. (notas de aula). Rio de Janeiro: PUC-RJ.

Leite, F. C. L. (2009). *Como gerenciar e ampliar a visibilidade da informação científica brasileira: repositórios institucionais de acesso aberto*. Retrieved from https://books.google.com.br/books?hl=pt-BR&lr=&id=Q6CyB8PhRH4C&oi=fnd&pg=PA94&dq=fernando+leite+unb+acesso+aberto&ots=-wfv0Ay5yD&sig=fOmITIM9ZqzAmIPyRpD2WMYkZD8.

Lin, J., & Fushman, D. D. (2005). Representation of information needs and the elements of context: A case study in the domain of clinical medicine. In *ACM SIGIR 2005 Workshop on Information Retrieval in Context (IRiX)*. Retrieved from http://www.umiacs.umd.edu/~jimmylin/publications/Lin_Demner-Fushman_SIGIR2005_workshop.pdf.

Liverman, C. T., Ingalls, C. E., Fulco, C. E., & Kipen, H. M. (1997). Understanding the information needs of health professionals. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK45472/

Lynch, C. A. (2003). Institutional repositories: essential infrastructure for scholarship in the digital age. *Portal: Libraries and the Academy*, 3(2), 327–336. https://doi.org/10.1353/pla.2003.0039.

Machado, C. M., Rebholz-Schuhmann, D., Freitas, A. T., & Couto, F. M. (2015). The semantic web in translational medicine: current applications and future directions. *Briefings in Bioinformatics*, 16(1), 89–103.

MacMullen, J. (2003). Discovery support systems and the integration of biomedical knowledge. *Canadian Journal of Information and Library Science*, 27(4), 78–79.

MacMullen, W. J., & Denn, S. O. (2005). Information problems in molecular biology and bioinformatics. *Journal of the American Society for Information Science and Technology*, 56(5), 447–456.

Maggio, L. A., Steinberg, R. M., Moorhead, L., O'Brien, B., & Willinsky, J. (2013). Access of primary and secondary literature by health personnel in an academic health center: implications for open access. *Journal of the Medical Library Association : JMLA*, 101(3), 205–212. https://doi.org/10.3163/1536-5050.101.3.010.

Malta, M. C., & Baptista, A. A. (2017). The development process of a metadata application profile for the social and solidarity economy. In Malta, M. C., Baptista, Ana Alice, & Walk, P. (Editors). *Developing metadata application profiles* (pp. 65–97). Hershey: IGI Global.

Malta, M., & Baptista, A. A. (2012). State of the art on methodologies for the development of a metadata application profile. MTSR 2012: Research Conference on Metadata and Semantics Research, pp 61-73. Retrieved from http://repositorium.sdum.uminho.pt/handle/1822/20490.

Malta, M. C., & Baptista, A. A. (2013). Me4DCAP V0. 1: a method for the development of Dublin Core Application Profiles. *Information Services & Use*, 33(2):161-171.

Malta, M. C. (2014). *Contributo metodológico para o desenvolvimento de perfis de aplicação no contexto da Web Semântica* (Tese de Doutoramento). Universidade do Minho.

Manfreda, K. L., Bosnjak, M., Berzelak, J., Haas, I., Vehovar, V., & Berzelak, N. (2008). Web *survey*s versus other *survey* modes: A meta-analysis comparing response rates. *Journal of the Market Research Society*, 50(1), 79.

Manola, F., & Miller, E. (2004). RDF Primer. Retrieved June 20, 2013, from http://www.w3.org/TR/2004/REC-rdf-primer-20040210/.

Manouselis, N., Najjar, J., Kastrantas, K., Salokhe, G., Stracke, C. M., & Duval, E. (2010). Metadata interoperability in agricultural learning repositories: an analysis. *Computers and Electronics in Agriculture*, 70(2), 302–320.

March, S. T., & Smith, G. F. (1995). Design and natural science research on information technology. *Decision Support Systems*, 15(4), 251–266.

Marcondes, C. H. (2011). Um modelo semântico de publicações eletrônicas: A semantic model for electronic publishing. *Liinc em Revista*, 7(1). Retrieved from http://revista.ibict.br/liinc/index.php/liinc/article/view/404.

Marcondes, C. H., & Sayão, L. F. (2009). À guisa de introdução: repositórios institucionais e livre acesso. *Implantação* e gestão de repositórios institucionais: políticas, memória, livre acesso e preservação. Salvador: EDUFBA, 9–21.

Markey, K. (2007a). Twenty-five years of end-user searching, Part 1: Research findings. *Journal of the American Society for Information Science and Technology*, 58(8), 1071–1081.

Markey, K. (2007b). Twenty-five years of end-user searching, Part 2: Future research directions. *Journal of the American Society for Information Science and Technology*, 58(8), 1123–1130. https://doi.org/10.1002/asi.20601

McCarthy, L. H., & Reilly, K. E. (2000). How to write a case report. *Family Medicine-Kansas City*, 32(3), 190–195.

McGuinness, D. L. & Van Harmelen, F. (2004). OWL web ontology language overview. *W3C Recommendation*, 10(10), 2004.

McKibbon, K., & Fridsma, D. B. (2006). Effectiveness of clinician-selected electronic information resources for answering primary care physicians' information needs. *Journal of the American Medical Informatics Association*, 13(6), 653–659.

Medicina. (2015, June 20). In *Wikipédia*, *a enciclopédia livre*. Retrieved from https://pt.wikipedia.org/w/index.php?title=Medicina&oldid=42655820.

Menon, D., & Stafinski, T. (2005). Bridging the "know-do" gap in healthcare priority-setting: what role has academic research played? In *Healthcare Management Forum* (Vol. 18, pp. 26–32). SAGE Publications Sage CA: Los Angeles, CA. Retrieved from http://journals.sagepub.com/doi/abs/10.1016/S0840-4704(10)60066-X.

Menzel, H. (1964). The information needs of current scientific research. *The Library Quarterly*, 34(1), 4–19.

Metadata. (2016, December 31). In *Wikipedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Metadata&oldid=757616273.

Miller, P. (2000). Interoperability: What is it and why should I want it? *Ariadne*, (24). Retrieved from http://www.ariadne.ac.uk/issue24/interoperability?ref=SevSevil.Com.

Miller, P. (2008, February 26). Sir Tim Berners-Lee: Semantic Web is open for business. Retrieved January 18, 2017, from http://www.zdnet.com/article/sir-tim-berners-lee-semantic-web-is-open-for-business/.

Miranda, S. V., & Tarapanoff, K. (2008). Information needs and information competencies: a case study of the off-site supervision of financial institutions in Brazil. *Information Research*, 13(2), 5.

Mons, B., & Velterop, J. (2009). Nano-publication in the e-science era. In *Workshop on Semantic Web Applications in Scientific Discourse (SWASD 2009)*. Retrieved from http://www.surf.nl/sitecollectiondocuments/nano-publication%20-%20mons%20-%20velterop.pdf.

Moresi, E. (2003). *Metodologia da pesquisa*. Brasília: UnB.

Mueller, S. P. M. (2000a). A ciência, o sistema de comunicação científica e a literatura científica. In B. S. Campello, B. V. Cendón, & Kremer (Eds.). *Fontes de informação para pesquisadores e profissionais* (pp. 21–34). Belo Horizonte: UFMG.

Mueller, S. P. M. (2000b). O periódico científico. In B. S. Campello, B. V. Cendón, & J. M. Kremer (Eds.). *Fontes de informação para pesquisadores e profissionais* (pp. 73–96). Belo Horizonte: UFMG.

Mueller, S. P. M. (2006). A comunicação científica e o movimento de acesso livre ao conhecimento. *Ciência Da Informação*, *Brasília*, 35(2), 27–38.

Murry Jr, J. W., & Hammons, J. O. (1995). Delphi: a versatile methodology for conducting qualitative research. *The Review of Higher Education*, *18*(4), 423–436.

Myers, M. D. (2013). *Qualitative research in business and management*. Sage Publications.

Myers, M. D., & Avison, D. (1997). Qualitative research in information systems. *Management Information Systems Quarterly*, 21, 241–242.

Nagamori, M., & Sugimoto, S. (2006). A metadata schema registry as a tool to enhance metadata interoperability. *Mitsuharu Nagamori and Shigeo Sugimoto*, 3(1). Retrieved from http://www.citeulike.org/group/318/article/883849.

National Cancer Institute. NCI Wiki. *Index of terminologies.* Available https://wiki.nci.nih.gov/display/VKC/ Index+of+Terminologies

National Information Standards Organization. (2004). *Understanding metadata*. Bathesda: Niso.

National Information Standards Organization. (2007). *A framework of guidance for building good digital collections*. Bathesda: NISO.

Neches R, Fikes R, Finin T, Gruber T, Petel R, Senator T, Swartout WR. Enabling technology for knowledge sharing. *AI Magazine,* v. 12, n. 3, p. 37-56, 1991.

Neufeld, M. L., & Cornog, M. (1986). Database history: From dinosaurs to compact discs. *Journal of the American Society for Information Science*, 37(4), 183.

Nilsson, M., Baker, T., & Johnston, P. (2009). Interoperability levels for Dublin Core metadata. Retrieved from http://dublincore.org/documents/interoperability-levels/.

Nilsson, M., Baker, T., & Johnston, P. (2008, January 14). *The Singapore Framework for Dublin Core Application Profiles*. Retrieved from http://dublincore.org/documents/singapore-framework/.

Nilsson, M., Miles, A. J., Johnston, P., & Enoksson, F. (2009). Formalizing Dublin Core Application Profiles: Description Set Profiles and Graph Constraints. In M.A. Sicilia & M. D. Lytras (Eds.). *Metadata and Semantics* (pp. 101–111). Springer US. Retrieved from http://*link*.springer.com/chapter/10.1007/978-0-387-77745-0_10.

Nunamaker Jr, J. F., & Chen, M. (1990). Systems development in information systems research. In *System Sciences*, *1990. Proceedings of the Twenty-Third Annual Hawaii International Conference on* (Vol. 3, pp. 631–640). IEEE. Retrieved from http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=205401.

Offermann, P., Levina, O., Schönherr, M., & Bub, U. (2009). Outline of a design science research process. In *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology* (p. 7). ACM. Retrieved from http://dl.acm.org/citation.cfm?id=1555629.

Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: an example, design considerations and applications. *Information & Mangement*, *42*(1), 15–29.

Osheroff, J. A., & Bankowitz, R. A. (1993). Physicians' use of computer software in answering clinical questions. *Bulletin of the Medical Library Association*, 81(1), 11.

Osheroff, J. A., Forsythe, D. E., Buchanan, B. G., Bankowitz, R. A., Blumenfeld, B. H., & Miller, R. A. (1991). Physicians' information needs: analysis of questions posed during clinical teaching. *Annals of Internal Medicine*, 114(7), 576–581. https://doi.org/10.7326/0003-4819-114-7-576.

Owen, P. (1995). Clinical practice and medical research: bridging the divide between the two cultures. *Br J Gen Pract*, 45(399), 557–560.

Packer, A. L., Santos, S., & Meneghini, R. (2017). SciELO preprints on the way. Retrieved from http://blog.scielo.org/en/2017/02/22/scielo-preprints-on-the-way/.

Paisley, W. J. (1968). Information needs and uses. *Annual Review of Information Science and Technology*, 3(1), 1–30.

Palmer, C. L., & Neumann, L. J. (2002). The information work of interdisciplinary humanities scholars: Exploration and translation. *The Library Quarterly*, 85–117.

Park, J., & Ram, S. (2004). Information systems interoperability: What lies beneath? *ACM Transactions on Information Systems (TOIS)*, 22(4), 595–632.

Park, J.-R., & Tosaka, Y. (2010). Metadata creation practices in digital repositories and collections: schemata, selection criteria, and interoperability. *Information Technology and Libraries*, 29(3), 104–116.

Passos, J., Sequeira, C., & Fernandes, L. (2017). Desenvolvimento de um catálogo CIPE: necessidades do idoso em enfermagem de saúde mental e psiquiatria. *Revista Portuguesa de Enfermagem de Saúde Mental*, (10), 15–23.

Patel, M., Koch, T., Doerr, M., & Tsinaraki, C. (2005). D5.3.1: Semantic Interoperability in Digital Library Systems. UKOLN. Retrieved from http://opus.bath.ac.uk/23606/1/SI_in_DLs.pdf.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems*, 24(3), 45–77.

Pereira, A. M., Ribeiro Jr, D. I., & Neves, G. L. C. (2006). Metadados para a descrição de recursos da Internet: as novas tecnologias desenvolvidas para o padrão Dublin Core e sua utilização. *Pesquisa Brasileira em Ciência da Informação e Biblioteconomia*, 1(1). Retrieved from http://periodicos.ufpb.br/index.php/pbcib/article/view/8872.

Peset, F., Ferrer-Sapena, A., & Subirats-Coll, I. (2011, March). Open data y Linked Open Data: su impacto en el área de bibliotecas y documentación. *El profesional de la información*, 20(2), 165-173. DOI: 10.3145/epi.2011.mar.06.

Powell, C. (2003). The Delphi technique:myths and realities. *J Adv Nurs*, *41*(4), 376–82.

Powell, A., Nilsson, M., Naeve, A., & Johnston, P. (2007). DCMI Abstract Model. *MyScienceWork*. Retrieved from https://www.mysciencework.com/publication/show/ad8c6acd16205d8dd583bdeb7c8bc304.

Prakasan, P. M. (2013). Information needs and use of healthcare professionals: International perspective. *DESIDOC Journal of Library and Information Technology*, 33(6), 465–473.

Pries-Heje, J., & Pries-Heje, L. (2015, July). *DR 01 Design Science Research*. PPT, Universidade do Minho.

Pudovkin, A., & Garfield, E. (2002). Algorithmic procedure for finding semantically related journals. *Journal of the Association for Information Science and Technology*, 53(13), 1113–1119. https://doi.org/DOI: 10.1002/asi.10153.

Reitz, J. M. (2004). *Dictionary for library and information science*. Libraries Unlimited. Retrieved from https://books.google.com.br/books?hl=pt-BR&lr=&id=f9WH9soOBrUC&oi=fnd&pg=PR9&dq=reitz+2004&ots=pX4JJ7fivt&sig=NLuN_wYrZtTsUklxiAwysLNvhY8.

Reynolds, N., Diamantopoulos, A., & Schlegelmilch, B. (1993). Pretesting in questionnaire design: a review of the literature and suggestions for further research. *Journal of the Market Research Society*, 35(2), 171–183.

Rodriguez, M. C. Three options are optimal for multiple-choice itens: a meta-analisys of 80 years of research. Educational Meausrement: issues and practice, v. 24, n. 2, p. 3-13, 2005.

Rowlands, I. (2007). Electronic journals and user behavior: A review of recent research. *Library & Information Science Research*, 29(3), 369–396.

Rowlands, I., & Nicholas, D. (2008). Understanding Information behaviour: how do students and faculty find books? *The Journal of Academic Librarianship*, 34(1), 3–15. https://doi.org/10.1016/j.acalib.2007.11.005

Sandusky, R. J., & Tenopir, C. (2008). Finding and using journal-article components: Impacts of disaggregation on teaching and research practice. *Journal of the American Society for Information Science and Technology*, 59(6), 970–982. https://doi.org/10.1002/asi.20804

Saracevic, T. (1995). Interdisciplinary nature of information Science. *Ciência da Informação*, 24(1):1-9.

Schaffert, S., Gruber, A., & Westenthaler, R. (2005). A semantic wiki for collaborative knowledge formation. *Proceedings of the Semantics*. Retrieved from http://www.wastl.net/download/SemWikiForCollKnowForm_20060120.pdf

Schaffner, A. C. (1994). The future of scientific journals: lessons from the past. *Information Technology and Libraries*, 13(4), 239.

Selvage, M. Y., Wolfson, D., Zurek, B., & Kahan, E. (2006). Achieve semantic interoperability in a SOA. *IBM Inc*.

Semantics. (2013, June 14). In *Wikipedia*, *the free encyclopedia*. Retrieved from http://en.wikipedia.org/w/index.php?title=Semantics&oldid=558708512.

Semple, N. (2006). Digital repositories. Retrieved July 16, 2015, from http://www.dcc.ac.uk/resources/briefing-papers/introduction-curation/digital-repositories.

Serviço Nacional de Saúde. (2015). Rede Nacional de Especialidade Hospitalar e de Referenciação de Oncologia Médica. Serviço Nacional de Saúde. Retrieved from http://www2.portaldasaude.pt/portal/conteudos/a+saude+em+portugal/publicacoes/estudos/gtrelatorio+oncologiamedica.htm.

Sheth, A., Agrawal, S., Lathem, J., Oldham, N., Wingate, H., Yadav, P., & Gallagher, K. (2006). Active semantic electronic medical record. In *The Semantic Web-ISWC 2006* (pp. 913–926). Springer. Retrieved from http://*link*.springer.com/chapter/10.1007/11926078_66.

Shotton, D. (2009). Semantic publishing: the coming revolution in scientific journal publishing. *Learned Publishing*, 22(2), 85–94. https://doi.org/10.1087/2009202.

Shotton, D., Portwin, K., Klyne, G., & Miles, A. (2009). Adventures in semantic publishing: exemplar semantic enhancements of a research article. *PLoS Computational Biology*, 5(4), e1000361. https://doi.org/10.1371/journal.pcbi.1000361

Simon, H. A. (1996). *The sciences of the artificial*. MIT press. Retrieved from https://books.google.com.br/books?hl=pt-PT&lr=&id=k5Sr0nFw7psC&oi=fnd&pg=PR9&dq=simon+ %22artificial+science%22&ots=-v1JIICOGt&sig=xqf-TIaK6iCMIrSeyTz5vJsDe54.

Semântica. In: Dicionário infopédia da Língua Portuguesa com Acordo Ortográfico [em linha]. Porto: Porto Editora, 2003-2017. Available: https://www.infopedia.pt/dicionarios/lingua-portuguesa/semântica.

Sintaxe In: Dicionário infopédia da Língua Portuguesa com Acordo Ortográfico [em linha]. Porto: Porto Editora, 2003-2017. Available: https://www.infopedia.pt/dicionarios/lingua-portuguesa/sintaxe

Smith, R. (1996). What clinical information do doctors need? *Bmj*, 313(7064), 1062–1068.

Soares, D. de S. (2010). *Interoperabilidade entre sistemas de informação na Administração Pública*. Tese, Retrieved from https://repositorium.sdum.uminho.pt/handle/1822/10539.

Sobrinho-Simões, M. (2016). *O Cancro*. Lisboa: Fundaçao Francisco Manuel Dos Santos.

Søndergaard, T. F., Andersen, J., & Hjørland, B. (2003). Documents and the communication of scientific and scholarly information: Revising and updating the UNISIST model. *Proceedings of the American Society for Information Science and Technology*, 40(1), 516–516.

Souza, M. I. F., & Alves, M. das D. R. (2009). Representação descritiva e temática de recursos de informação no sistema agência Embrapa: uso do padrão dublin core Descriptive and thematic representation of electronic information of the Embrapa information agency: uses of the dublin core metadata standard. *RDBCI*, 7(1). Retrieved from http://www.sbu.unicamp.br/seer/ojs/index.php/rbci/article/view/430

Stedman, T. L. (2003). *Stedman dicionário médico*. Rio de Janeiro: Guanabara Koogan.

Stewart, B., Wild, C. P.(Ed.). (2014). World cancer report 2014. Retrieved from http://www.thehealthwell.info/search-results/world-cancer-report-2014?source=relatedblock.

Steyskal, S., & Coyle, K. (2017, February 9). SHACL use cases and requirements. Retrieved February 9, 2017, from https://www.w3.org/TR/shacl-ucr/.

Suber, P. (2007). Open access overview. Retrieved from Peter Suber's Website: Http://www.Earlham.Edu/~ Peters/fos/overview.Htm. Retrieved from http://legacy.earlham.edu/~peters/fos/overview.htm.

Stuart, D. (2016). Ontologies and the semantic web. In: *Practical Ontologies for Information Professionals*. London: Facet Publishing. pp. 27-51.

Swan, A. (2006). Overview of scholarly communication. Retrieved from http://eprints.soton.ac.uk/262427.

Swan, A. (2011). Institutional repositories-now and next. Retrieved from http://eprints.ecs.soton.ac.uk/21471/1/Chapter_8_Swan_postprint.doc.

Taurion, C. (2014). *O que é um modelo preditivo*? Retrieved April 14, 2017, from https://www.google.com.br/search?q=Cezar+Taurion+%22modelo+preditivo%22&oq=Cezar+Taurion+%22modelo+preditivo%22&aqs=chrome..69i57j69i60.1071j0j8&sourceid=chrome&ie=UTF-8.

Taylor, R. S. (1962). The process of asking questions. *American Documentation*, 13(4), 391–396.

Tedd, L. A., & Large, J. A. (2004). *Digital libraries:* principles and practice in a global environment. Retrieved from https://books.google.com.br/books?hl=pt-BR&lr=&id=u_Gi1CPuP3kC&oi=fnd&pg=PR11&dq=DIGITAL+LIBRARIES:+PRINCIPLES+AND+PRACTICE+IN+A+GLOBAL+ENVIRONMENT&ots=qGZ695mDyB&sig=oGUWBNHQSRbiwnxNaNHRuFXEiGU.

Tenopir, C., King, D. W., Edwards, S., & Wu, L. (2009b). Electronic journals and changes in scholarly article seeking and reading patterns. In *Aslib proceedings* (Vol. 61, pp. 5–32). Retrieved from http://www.emeraldinsight.com/journals.htm?articleid=1766871&show=abstract.

Tenopir, C., King, D. W., Spencer, J., & Wu, L. (2009a). Variations in article seeking and reading patterns of academics: What makes a difference? *Library & Information Science Research*, 31(3), 139–148.

The Netherlands Cancer Institute Antoni van Leeuwenhock Hospital. (2010). The Netherlands Cancer Institute Scientific Annual Report 2010.

Timpka, T., Ekström, M., & Bjurulf, P. (1989). Information needs and information seeking behaviour in primary health care. *Scandinavian Journal of Primary Health Care*, 7(2), 105–109.

Uniform resource identifier. (2013, June 19). In *Wikipedia*, *the free encyclopedia*. Retrieved from https://en.wikipedia.org/w/index.php?title=Uniform_resource_identifier&oldid=557279170.

United Kingdom Academy for Information Systems. (1997, December). Definition of Information Systems. *Newsletter for the UK Academy for Information Systems*. Manchester. Retrieved from http://www.scs.leeds.ac.uk/ukais/Newsletters/vol3no4.html#Definition

VanGeest, J. B., Johnson, T. P., & Welch, V. L. (2007). Methodologies for improving response rates in *survey*s of physicians a systematic review. *Evaluation & the Health Professions*, 30(4), 303–321.

Vaishnavi, V., Kuechler, W. Design research in information systems. Retrieved http://desrist.org/ desrist/content/design-science-research-in-information-systems.pdf.

Victorino, M. (2012). Organização da Informação para dar Suporte à Arquitetura Orientada a Serviços: Reuso da Informação nas Organizações. 2011. 276 p. Tese (Doutorado). Faculdade de Ciência da Informação, Programa de Pós-Graduação em Ciência da Informação, Universidade de Brasília-UnB. Brasília.

Vocabulário. In: Dicionário infopédia da Língua Portuguesa com Acordo Ortográfico [em linha]. Porto: Porto Editora, 2003-2017. Available: https://www.infopedia.pt/dicionarios/lingua-portuguesa/vocabulário

W3C. (2004). *RDF Vocabulary Description Language 1.0*: RDF Schema. Retrieved June 22, 2013, from http://www.w3.org/TR/rdf-schema/.

W3C. (2011). *OpenMetadataRegistry:* Semantic Web Standards. Retrieved May 28, 2016, from https://www.w3.org/2001/sw/wiki/OpenMetadataRegistry.

W3C. (2013). *SPARQL 1.1 Overview*. Retrieved June 20, 2013, from http://www.w3.org/TR/sparql11-overview/

Weibel, S. (1997). The Dublin core: a simple content description model for electronic resources. *Bul. Am. Soc. Inf. Sci*, 24(1), 9–11.

Weinberg, B. H. (1988). Why indexing fails the researcher. *The Indexer*, 16(1), 3–6.

White, H. C. (2010). Considering personal organization: metadata practices of scientists. *Journal of Library Metadata*, 10(2–3), 156–172. https://doi.org/10.1080/19386389.2010.506396

Wildemuth, B. M. (2016). Applications of Social Research Methods to Questions in Information and Library Science, 2nd Edition. ABC-CLIO.

Wilson, T. D. (1994). Information needs and uses: fifty years of progress. *Fifty Years of Information Progress: A Journal of Documentation Review*, 15–51.

Wilson, T. D. (2000). Human information behavior. *Informing Science*, *3*(2), 49–56.

Wilson, T. D. (1971). Research priorities in social welfare library and information work. *InformationR.net*. Retrieved from http://www.informationr.net/tdw/publ/papers/jlib75.html

Woodley, M. S., Clement, G., & Winn, P. (2005a). Metadata. Retrieved March 31, 2017, from http://dublincore.org/documents/usageguide/glossary.shtml

Woodley, M. S., Clement, G., & Winn, P. (2005c). Metadata schema registry. *DCMI Glossary*. Retrieved from http://dublincore.org/documents/usageguide/glossary.shtml

Woodley, M. S., Clement, G., & Winn, P. (2005b, November 7). Semantic interoperability. *DCMI Glossary*. Retrieved from http://dublincore.org/documents/usageguide/glossary.shtml#S

Woodley, M. S., Clement, G., & Winn, P. (2013). Semantic web. *DCMI Glossary*. Retrieved from http://dublincore.org/documents/usageguide/glossary.shtml

Woolf, S. H., & Benson, D. A. (1989). The medical information needs of internists and pediatricians at an academic medical center. *Bulletin of the Medical Library Association*, *77*(4), 372.

World Health Organization. (2016). Cancer. Retrieved March 17, 2017, from http://www.who.int/cancer/en/

Wyatt, J. C. (2000). Clinical questions and information needs. *Journal of the Royal Society of Medicine*, *93*(4), 168.

Yin, R. K. (2013). *Case Study Research: Design and Methods*. Thousand Oaks: SAGE Publications.

Young, H. (1988). *Glosario ALA de bibliotecología y ciencias de la información*. Retrieved from https://dialnet.unirioja.es/servlet/libro?codigo=302732

Zawawi, S., & Majid, S. (2001). The information needs and seeking behaviour of the IMR biomedical scientists. *Malaysian Journal of Library & Information Science*, 5(1), 25–41.

GLOSSÁRIO

Agente de *software* – um programa de computador que executa tarefas em nome de outra entidade. Freqüentemente usado para fazer referência a um programa que pesquisa na Internet informações que atendem aos requisitos especificados por um usuário individual.

Alinhamento de vocabulário – processo de análise múltipla de vocabulários para determinar quais termos são comuns entre eles e o registros aos quais estão relacionados (W3C, 2013).

Aplicação – um programa de aplicação (aplicativo) é um programa de computador projetado para executar um grupo de funções coordenadas, tarefas ou atividades em beneifício do usuário (Wikepedia, 2017)

Artefacto – é algo artificial ou construído por humanos (Simon, 1996)

Atributo –"dado significativo que pertence a uma classe, que geralmente contém valores que descrevem cada instância da classe" (Glossary of Unified Modelling..., 2017)

Autoarquivamento – é o ato de o autor depositar uma cópia de um documento eletrônico on-line acessível ao público (Wikepedia, 2017).

Bibliotecas digitais – "[...] conjunto de objetos digitais construídos a partir do uso de instrumentos eletrônicos, concebidos com o objetivo de registrar e comunicar pensamentos, ideias, imagens e sons, disponíveis a um contingente ilimitado de pessoas, dispersas onde quer que a plataforma WWW alcance" (Alvarenga, 2003), de forma que considerações em relação a esse tipo de biblioteca é extensivo aos Ris.

Categorização – é realizada por meio da análise do domínio a partir de recortes que permitem identificação dos conceitos (categorias) que fazem parte do domínio (Campos & Gomes, 2008).

Ciência da Informação – "[...] área dedicada à investigação científica e à prática profissional abordando os problemas de comunicação eficaz do conhecimento e dos registros de conhecimento entre os seres humanos no contexto social, institucional e ou das necessidades e uso da informação" (Saracevic, 1995, p. 2)

Classe – "é a representação de um conjunto de coisas reais ou abstratas que são reconhecidas como sendo do mesmo tipo por compartilhar as mesmas características, atributos, relações e semântica" (Furlan, 1998)

Coleção – refere-se a um número de documentos (livros, relatórios e registros, etc) reunidos numa localização física ou virtual única por uma ou por mais pessoas, ou por uma entidade corporativa organizada de forma sistemática com vista a facilitar a recuperação (Reitz, 2004).

Comportamento informacional – "[...] a totalidade do comportamento humano em relação ao uso de fontes e canais de informação, incluindo a busca da informação passiva ou ativa, e uso da informação" (Wilson, 2000, p. 39)

Comportamento do uso de informação – consiste de umo conjunto dos atos físicos e mentais na incorporação da informação encontrada em conhecimentos existentes no indivíduo (Wilson, 2000, p. 39).

Comportamento de busca de informação – É a busca intencional de informação, em consequência da necessidade de satisfazer algum objetivo (Wilson, 2000, p. 39).

Creative Commons –"são várias licenças públicas que permitem a distribuição gratuita de uma obra protegida por direitos autorais" (Wikepedia, 2017)

Desenho de investigação – "conjunto de técnicas analíticas e perspectivas (complementando as perspectivas positivista e interpretativista) para realizar investigação em sistemas de informação" (Vaishnavi & Kuechler, 2004)

Digital Object Identifier (DOI) – sistema para identificação e intercâmbio da propriedade intelectual no ambiente digital.

DBpedia – é um esforço da comunidade para extrair informações estruturadas da Wikipedia e torná-las disponíveis na Web. DBpedia é muitas vezes descrita como um hub para a nuvem de dados. Uma representação RDF dos metadados contidos na Wikipédia e disponibilizada para a consulta SPARQL na World Wide Web (W3C, 2013).

Dublin Core Metadata Element Set – vocabulário de 15 propriedades para uso na descrição de recursos. O conjunto completo de vocabulários, DCMI Metadata Terms, também inclui conjunto de classes, esquemas de codificação de vocabulário, esquemas de codificação de sintaxe ("Dublin Core Metadata Element Set, Version 1.1," 2012)

Elementos de metadados – propriedade de um objeto que é definida em uma especificação. "autor/creator", título/"title" e autor/"subject" são propriedades definidas como elementos de metadados. Definido também como sinônimo de propriedade e de atributo (Powell et al., 2007).

Environmental Scan **("mapeamento de domínio")** – "[...] é um relatório que contém uma revisão dos esquemas de metadados que estão disponíveis em qualquer serialização da web semântica (RDF/XML e Turtle) e que poderão servir as necessidades do modelo de domínio" (Malta & Baptista, 2017, p. 105)

Esquema de metadados – Conjunto de elementos de metadados designados para um objetivo específico, como a descrição de um tipo particular de recurso de informação (Chan & Zeng, 2006, National Information Standards Organization, 2004).

Esquemas de codificação – apresentam informação contextual ou interpretação de regras que auxiliam na interpretação do valor dos termos. Para tanto utiliza vocabulários controlados, notações formais ou regras de interpretação que possibilitam o entendimento por humanos, quando não forem interpretadas por um agente de software (Dublin Core Metadata Iniciative, 2012, p. 1). A DCMI considera dois esquemas de codificação: o esquema de sintaxe (SES) e o esquema de vocabulário (VES).

Granularidade de um recurso – refere-se ao tamanho, à decomponibilidade e ao grau em que um recurso é planejado para ser usado como parte de um recurso maior (NISO, 2007).

Harvester **de metadados**– é um aplicativo cliente que emite solicitações OAI-PMH. Um harvester de metadados é operado por um prestador de serviços como um meio de coletar metadados de repositórios (http://www.openarchives.org/OAI/openarchivesprotocol.html#harvester) para um catalogo central (NISO, 2004).

Hipertexto –escrita não sequencial. O termo foi cunhado por Ted Nelson para designar relações entre partes de informações. Hoje em dia inclui outras mídias além do texto e às vezes é denominado de hipermídia (Wikipedia).

Hypertext transfer protocol **(http)** – um protocolo de computador para transferir informação através da rede de forma a satisfazer as exigências de um sistema de hipertexto global (W3C, 2013).

Identificador – um identificador é um nome inequívoco para um recurso.

Indexação – o processo de avaliação de entidades de informação e criação de termos que auxiliam na descoberta e acesso à entidade. Os termos do índice podem ser em linguagem natural ou vocabulário controlado ou uma notação de classificação.

Inferência – é o processo de derivar conclusões lógicas a partir de um conjunto de pressupostos iniciais. Usando *Linked Data*, os relacionamentos existentes são modelados como um conjunto de relações (nomeadas) entre recursos. *Linked Data* ajuda os seres humanos e as máquinas a encontrarem novas relações através de procedimentos automáticos que geram novas relações baseadas nos dados e com base em alguma informação adicional na forma de um vocabulário (W3C, 2013).

Instância –uma ocorrência específica de um objeto que pertence a uma classe (Instance, 2017).

Interoperabilidade semântica – corresponde ao significado dos elementos de dados e os relacionamentos entre eles. Inclui o desenvolvimento de vocabulários para descrever intercâmbio de dados, e assegurar que os elementos de dados possuam o mesmo significado para as partes que se comunicam (European Commission, 2010, p. 23).

Interoperabilidade –interoperabilidade é a capacidade de múltiplos sistemas, utilizando diferentes *hardwares* e *softwares*, estruturas de dados, e interfaces, trocar e partilhar dados (NISO, 2004).

Investigação – "investigação original realizada para contribuir com o conhecimento e a compreensão em um campo particular "(Myers, 2013).

Link – expressa uma ou mais (explicito ou implicita) relação entre dois ou mais recursos (W3C, 2017).

Linked Data – Um padrão para hiperligar conjuntos de dados legíveis por máquina uns aos outros usando tecnologias da Web Semântica, especialmente através do uso de RDF e URIs. Permite consultas SPARQL distribuídas de conjuntos de dados e a navegação ou descoberta para encontrar informações (quando comparada a estratégia de pesquisa). *Linked Data* destina-se ao acesso por seres humanos e máquinas (W3C, 2013).

Linked Open Data – é a publicação *Linked Data* sobre licença aberta, permitindo seu reuso livremente (W3C, 2013).

Literal – é definido como "[...] uma entidade que utiliza uma sequência de caracteres Unicode como uma forma lexical, juntamente com uma tag de idioma opcional ou tipo de dados, para denotar um recurso [...]. Um literal oferece uma grande quantidade de flexibilidade e poder, mas aumenta a complexidade. Os metadados devem também incluir um literal apropriado que reflita o valor base da entidade de metadados. Por exemplo, nesses fragmentos: criador = "Público, John Q." Creator = "http://authority.org/public-john-q-1234" o primeiro tem um valor expresso como um literal apropriado enquanto que o segundo tem um *link* (hipotético) para uma estrutura de autoridade. Não está inteiramente claro o que uma pessoa ou aplicação encontrará no final do link, então os metadados devem conter um literal apropriado para fins de descoberta simples.

Mapeamento de metadados (*Crosswalk*) – "[...] é o mapeamento de elementos, semântica e sintaxe de um esquema de metadados para outro" (NISO, 2004, p. 11)

Metadado descritivo – metadado que possibilita a descoberta de um objeto.

Metadados – são definidos como "dados sobre dados" ou "dados estruturados sobre dados". Os metadados incluem dados associados com qualquer sistema de informação ou objecto de informação com os seguintes objectivos: descrição, administração, requisitos legais, funcionalidades técnicas, uso e preservação (Woodley, Clement & Winn, 2005).

Método de investigação – "[...] estratégia de investigação que se move dos pressupostos filosóficos subjacentes para o projeto de pesquisa e coleta de dados" (Myers & Avison, 1997).

Modelo de dados – [...] é um modelo abstrato que organiza elementos de dados e padroniza como eles se relacionam entre si e com propriedades das entidades do mundo real (Wikepedia, 2017).

Modelo de domínio – é uma descrição de coisas que os metadados descreverá, e as relações entre essas coisas (Coyle & Baker, 2009).

Modelos preditivos - "É uma função matemática que, aplicada a uma massa de dados, consegue identificar padrões ocultos e prever o que poderá ocorrer" (Taurion, 2014).

Namespace – é uma forma de identificar inequivocamente (URI) o uso específico de um elemento de metadado para o esquema onde a definição é encontrada (NISO, 2004).

Não literal - é um valor que é físico, digital ou conceitual (Powell, et al., 2007).

Necessidades de informação - estado ou processo iniciado quando alguém percebe que existe uma lacuna entre a informação e o conhecimento disponíveis para resolver um problema e a sua solução" (Miranda & Tarapanoff, 2008).

OAI-ORE - especificações que permitem que repositórios distribuídos compartilhem informações sobre agregações de recursos da Web.

Objeto – sequência de *bits* digital (DRIVER, 2009). No contexto RDF, o objeto é a parte final de uma declaração RDF (W3C 2013).

Objeto digital – no contexto dos arquivos e bibliotecas digitais, é um registro de informação codificado digitalmente, formado de conteúdo informacional, metadados e identificador.

Ontologia - especificação formal, explícita de uma conceitualização compartilhada" (Gruber, 1993).

Open Archives Initiative Protocol For Metadada Harvesting (OAI-PMH) – é um protocolo desenvolvido pela Open Archives Initiative que define um mecanismo para disponibilização e coleta (*harvesting*) de registros de metadados em repositórios e proporciona a interoperabilidade.

OWL – linguagem de marcação de ontologias para a Web Semântica. Ela é baseada nas linguagens OIL e DALM-OIL e é uma recomendação da W3C (McGuinness & Van Harmelen, 2004).

Perfil de aplicação de metadados – "é um documento ou conjunto de documentos que especifica e descreve os metadados utilizados em uma aplicação em particular". Possui os seguintes componentes: requisitos funcionais, modelo de domínio, perfil de um conjunto de descrições, e os guias de uso, syntax e formato de dados" (Nilsson, Baker, & Johnston, 2008, p. 1)

Perfil do utilizador - "[...] representação estruturada das necessidades dos utilizadores em que um sistema de recuperação deveria se basear para responder às questões desses utilizadores" (Amato & Stracci, 1999).

Pesquisa documental - "é constituída pelo exame de materiais que ainda não receberam um tratamento analítico ou que podem ser reexaminados com vistas a uma interpretação nova ou complementar".

Política de acesso – declaração formal escrita pelos membros responsáveis pelo gerenciamento de arquivos, coleções especiais e materiais específicos disponíveis para o acesso, incluindo as condições e restrições de uso (Reitz, 2004).

Ponto de acesso – nome, terminologia, código por onde se pode buscar e identificar um registro bibliográfico. Em sistemas de busca e armazenamento de informação por computador, campo que permite ter acesso a um registro ou arquivo (Young, 1988).

Precisão – é a "relação entre o número de documentos pertinentes recuperados e o número total de documentos recuperados" (Meadows, 1999).

Preprint - Trabalho aceito para publicação em um periódico após revisão por pares, mas que ainda não foi publicado (Packer, Santos, & Meneghini, 2017)

Propriedade – é um aspecto específico, característica, atributo ou relação utilizada para descrever um recurso. Ainda pode ser denominada de campo, atributo e predicado. Ex.: Os elementos do DC são propriedades (DCMI, 2005).

Publicações semânticas – é o aprimoramento das publicações acadêmicas pelo uso de padrões Web para melhorar a interatividade, a abertura e a usabilidade, incluindo o uso de ontologias para codificar a semântica sob a forma de metadados em RDF legíveis por máquina (Shotton, 2013).

Recurso – no contexto RDF, um recurso pode ser qualquer coisa descrita por um gráfico RDF. Um recurso pode possuir um identificador de recursos internacionalizado (IRI).

Registro de metadados – um sistema publicamente acessível que grava a semântica, a estrutura e os formatos de intercâmbio de qualquer tipo de metadados. Uma autoridade formal, ou agência, mantém e gerencia o desenvolvimento ea evolução de um registro de metadados. A autoridade é responsável pelas políticas relativas ao conteúdo do registro e operação. Ver também http://www.dlib.org/dlib/may02/wagner/05wagner.html.

Repositório Digital – "sistema em rede que fornece serviços pertencentes a uma coleção de objetos digitais" (Bekaert & Van De Sompel, 2006).

Resource Description Framework (RDF) – é a linguagem básica para escrever metadados; Uma fundação que fornece uma arquitetura robusta e flexível para o processamento de metadados na Internet. O RDF manterá a capacidade de trocar metadados entre as comunidades de aplicativos, permitindo que cada comunidade defina e use os metadados que melhor atendam às suas necessidades.

Semântica – significado das palavras, por oposição à sua forma (Dicionário Infopedia..., 2017).

Serviço de descoberta – um serviço de descoberta é um serviço que permite que os agentes recuperem a descrição de recursos relacionados a serviços da Web.

Simple knowledge organization systems (SKOS) – fornece um enquadramento básico para a construção de esquemas de conceitos, que podem ser definidos como conjuntos de conceitos, opcionalmente incluindo declarações sobre relacionamentos entre estes conceitos (Catarino & Baptista, 2008)

Sintaxe – conjunto de regras sintáticas de uma dada época, autor etc. (Dicionário Infopedia..., 2017).

Sistema de informação – Conjuntos de componentes relacionados entre si na coleta, processamento, armazenamento, análise, disseminação e disponibilização (web) de informações, em determinada área, problema ou missão, com o propósito específico de atender usuários interessados (Pinheiro & Ferrez, 2014)

Sistemas de Informação (domínio) - é uma área multidisciplinar que aborda uma gama de atividades estratégicas, gerenciais e operacionais que envolvem a coleta, o processamento, o armazenamento, a distribuição e o uso de informações, e suas Tecnologias associadas, em sociedade e organizações (UK Academy of Information Systems apud Myers & Avison , 1997).

Taxonomia – pode referir-se tanto a uma classificação hierárquica das coisas como aos princípios que baseiam a classificação (Hodje, 2000).

Técnica de coleta de dados – o conjunto de processos e instrumentos elaborados para garantir o registro das informações, o controle e a análise dos dados" (Moresi, 2003).

Tecnologias semânticas – conjunto de tecnologias relacionadas a extração, representação, armazenamento, recuperação e análise de informações legíveis por máquina (W3C, 2013).

Termo – "um termo é uma propriedade (elemento), classe, esquema de codificação de vocabulário ou esquema de codificação de sintaxe" (Powell et al., 2017).

Triplo – é uma declaração RDF, que consiste de duas coisas (um sujeito e um objeto) e uma relação entre um verbo ou predicado. A forma do triplo RDF é sujeito-predicato-objeto (W3C, 2013).

Uniform Resource Identifier (URI) – identificador global no contexto da World Wide Web.

Vocabulário – "conjunto de termos especializados numa ciência ou arte; terminologia" (Dicionário Infopédia, 2017).

Vocabulário Controlado – é uma lista predeterminada de termos em determinado domínio (NISO, 2017).

Web Semântica – evolução ou parte da World Wide Web que consiste em dados legíveis por máquina em RDF e uma capacidade de consultar essas informações em formatos padronizados (W3C, 2013).



Application Profiles and Metadata Schemes

Scholarly Digital Repositories

I am a doctoral student at the University of Minho and I am in the process of determining the current praxis on the use of application profiles and metadata schemes in the scope of scholarly digital repositories. With this in mind, I would like to ask your kind contribution by answering a very small questionnaire. It will take you between one and five minutes to complete.

If you do not feel comfortable with these themes, we kindly ask you to forward this form to one of your colleagues who you think might be in position to contribute.

Thank you in advance,

Morgana Andrade

What is your Repository name and acronym?

What type of Repository is it?

- [ ] Institutional
- [ ] Thematic
- [ ] Organizational
- [ ] Learning object
- [ ] E-theses

Other (please specify)

What kind of material is stored on repository?

- [ ] Books / book chapters
- [ ] Reviews of books
- [ ] Papers presented at conferences
- [ ] Journal articles
- [ ] Journal articles submitted (pre-print)
- [ ] Dissertation or thesis
- [ ] Graduation projects
- [ ] Learning objects
- [ ] Patents
- [ ] Reports
- [ ] Research data
- [ ] Documents for discussion and work
- [ ] Historic documents
- [ ] Audiovisual materials
- [ ] Softwares

Other (please specify)



Application Profiles and Metadata Schemes

APPLICATION PROFILE

“An Application Profile (AP) is a declaration of which metadata terms an organization, information resource, application, or user community uses in its metadata” (Baker (2003). DCMI Usage board review of application profiles. Retrieved from: http://dublincore.org/usage/documents/profiles/index.shtml)

With which of the following international specifications is your Repository compliant (please select all that apply)?:

- [ ] Dublin Core Metadata Element Set (DCMI recommendation)
- [ ] DCMI Metadata Terms (DCMI Recommendation)
- [ ] Guidelines for Dublin Core Application Profiles (DCMI Recommendation)
- [ ] W3C RDF suite of specifications
- [ ] W3C OWL suite of specifications
- [ ] W3C SKOS suite of specifications
- [ ] Open Access Initiative-Protocol Metadata Harvesting (OAI-PMH)
- [ ] Open Archives Initiative Object Reuse and Exchange (OAI-ORE) Standard

Other (please specify)

Does your Repository adopt a specific application profile?

- ( ) Yes
- ( ) No

If you answered YES, please indicate the Web page or URI of Application Profile



Application Profiles and Metadata Schemes

METADATA SCHEME

A metadata scheme consists of a set of elements designed for a specific purpose, such as describing a particular type of information resource (NISO (2004). Understanding metadata. Bathesda: NISO. Retrieved from: http://www.niso.org/publications/press/UnderstandingMetadata.pdf)

Please select the ones used in your Repository

- [ ] Academic Metadata Format (AMF)
- [ ] Dublin Core Metadata Element Set (DC)
- [ ] DIM
- [ ] Electronic Thesis and Dissertation Metadata Standard (ETDMS)
- [ ] Learning Object Metadata (LOM)
- [ ] Machine-Readable Cataloguing (MARC)
- [ ] Metadata Encoding and Transmission Standard (METS)
- [ ] Metadata Object Description Schema (MODS)
- [ ] Moving Picture Experts Group (MPEG-21)
- [ ] Open Archives Initiative-Dublinc Core (OAI-DC)
- [ ] Open Digital Rights Language (ODRL)
- [ ] Metadata for exchanging business cards (vCard)
- [ ] Other (please specify)

A metadata element is “a property of a resource. As intended here, “properties” are attributes of resources characteristics of a resource, such as a Title, Publisher, or Subject. Elements are formally defined terms which are used to describe attributes and properties of a resources”! (Woodlley, M. S. (2005). DCMI glossary. Retrieved from: http://dublincore.org/documents/usageguide/glossary.shtml).

Please select the metadata elements used by your Repository (selecta all that apply). If this is not used, proceed to the next.

DUBLIN CORE (http://dublincore.org/documents/usageguide/elements.shtml)

- [ ] Title
- [ ] Subject
- [ ] Description
- [ ] Type
- [ ] Source
- [ ] Relation
- [ ] Coverage
- [ ] Creator
- [ ] Publischer
- [ ] Contributor
- [ ] Rights
- [ ] Date
- [ ] Format
- [ ] Identifier
- [ ] Language

Other (please specify)

LOM (http://ltsc.ieee.org/rdf/lomv1p0/terms#). If this is not used, proceed to the next.

- [ ] General -> identifier -> catalog
- [ ] General -> identifier -> entry
- [ ] General -> title
- [ ] General ->language
- [ ] General -> description
- [ ] General -> keyword
- [ ] General -> coverage
- [ ] General -> structure
- [ ] General -> aggregation level

- [ ] Life Cycle -> version
- [ ] Life Cycle -> status
- [ ] Life Cycle -> contribute -> role
- [ ] Life Cycle -> contribute -> entitity
- [ ] Life Cycle -> contribute -> date
- [ ] Meta-metadata -> Identifier -> catalog
- [ ] Meta-metadata -> identifier -> entry
- [ ] Meta-metadata -> contribute -> role
- [ ] Meta-metadata -> contribute -> entity
- [ ] Meta-metadata -> contribute -> date
- [ ] Meta-metadata -> metadata schema
- [ ] Meta-metadata -> language
- [ ] Technical -> format
- [ ] Technical -> size
- [ ] Technical -> location
- [ ] Technical -> requirement -> OrComposite -> type
- [ ] Technical -> requirement -> OrComposite -> name
- [ ] Technical -> requirement -> OrComposite -> minimum version
- [ ] Technical -> requirement -> OrComposite -> maximum version
- [ ] Technical -> installation remarks
- [ ] Technical -> other plataform requirements
- [ ] Technical -> duration
- [ ] Educational -> interactivivy type
- [ ] Educational -> learning resource type
- [ ] Educational -> interactivity level
- [ ] Educational -> semantic density
- [ ] Educational -> intended end user role
- [ ] Educational -> context
- [ ] Educational -> typical age range
- [ ] Educational -> difficulty
- [ ] Educational -> typical learning time
- [ ] Educational -> description
- [ ] Educational -> language
- [ ] Rights -> cost

- [ ] Rights -> copyright and other restrictions
- [ ] Rights -> description
- [ ] Relation -> kind
- [ ] Relation -> resource -> identifier -> catalog
- [ ] Relation -> resource -> identifier -> entry
- [ ] Relation -> resource -> description
- [ ] Annotation -> entitity
- [ ] Annotation -> date
- [ ] Annotation -> description
- [ ] Classification -> purpose
- [ ] Classification -> taxon path -> source
- [ ] Classification -> taxon path -> taxon -> id
- [ ] Classification -> taxon path -> taxon -> entry
- [ ] Classification -> description
- [ ] Classification -> keyword

Other (please specify)

MODS (http://www.loc.gov/standards/mods/mods-outline-3-5.html#top_level). If this is not used, proceed to the next.

- [ ] titleInfo -> title
- [ ] titleInfo -> subTitle
- [ ] titleInfo ->partNumber
- [ ] titleInfo -> partName
- [ ] titleInfo -> nonSort
- [ ] name -> namePart
- [ ] name -> displayForm
- [ ] name -> affiliation
- [ ] name -> role
- [ ] name -> description
- [ ] name -> etal
- [ ] typeOfResource
- [ ] genre
- [ ] originInfo -> place
- [ ] originInfo -> publisher

- [ ] originInfo -> dateIssued
- [ ] originInfo -> dateCreated
- [ ] originInfo -> dateCaptured
- [ ] originInfo -> dateValid
- [ ] originInfo -> dateModified
- [ ] originInfo -> copyrightDate
- [ ] originInfo -> dateOther
- [ ] originInfo -> edition
- [ ] originInfo -> issuance
- [ ] originInfo -> frequency
- [ ] language -> languageTerm
- [ ] language -> scriptTerm
- [ ] physicalDescription -> form
- [ ] physicalDescription -> reformattingQuality
- [ ] physicalDescription -> internetMediaType
- [ ] physicalDescription -> extent
- [ ] physicalDescription -> digitalOrigin
- [ ] physicalDescription -> note
- [ ] abstract
- [ ] tableOfContents
- [ ] targetAudience
- [ ] note
- [ ] subject -> topic
- [ ] subject -> geographic
- [ ] subject -> temporal
- [ ] subject -> titleInfo -> title
- [ ] subject -> titleInfo -> subTitle
- [ ] subject -> titleInfo -> partNumber
- [ ] subject -> titleInfo -> partName
- [ ] subject -> titleInfo -> nonSort
- [ ] subject -> name -> namePart
- [ ] subject -> name -> displayForm
- [ ] subject -> name -> affiliation
- [ ] subject -> name -> role

- [ ] subject -> name -> description
- [ ] subject -> genre
- [ ] subject -> hierarchicalGeographic -> continent
- [ ] subject -> hierarchicalGeographic -> country
- [ ] subject -> hierarchicalGeographic -> province
- [ ] subject -> hierarchicalGeographic -> region
- [ ] subject -> hierarchicalGeographic -> state
- [ ] subject -> hierarchicalGeographic -> territory
- [ ] subject -> hierarchicalGeographic -> county
- [ ] subject -> hierarchicalGeographic -> city
- [ ] subject -> hierarchicalGeographic -> island
- [ ] subject -> hierarchicalGeographic -> area
- [ ] subject -> hierarchicalGeographic -> extraterrestrialArea
- [ ] subject -> hierarchicalGeographic -> citySection
- [ ] subject -> cartographics
- [ ] subject -> geographicCode
- [ ] subject -> occupation
- [ ] classification
- [ ] relatedItem -> titleInfo
- [ ] relatedItem -> name
- [ ] relatedItem -> typeOfResource
- [ ] relatedItem -> genre
- [ ] relatedItem -> originInfo
- [ ] relatedItem -> language
- [ ] relatedItem -> physicalDescription
- [ ] relatedItem -> abstract
- [ ] relatedItem -> tableOfContents
- [ ] relatedItem -> targetAudience
- [ ] relatedItem -> note
- [ ] relatedItem -> subject
- [ ] relatedItem -> classification
- [ ] relatedItem -> relatedItem
- [ ] relatedItem -> identifier
- [ ] relatedItem -> location

- [ ] relatedItem -> accessCondition
- [ ] relatedItem -> part
- [ ] relatedItem -> extension
- [ ] relatedItem -> recordInfo
- [ ] identifier
- [ ] location -> physicalLocation
- [ ] location -> shelfLocator
- [ ] location -> url
- [ ] location -> holdingSimple
- [ ] location -> holdingExternal
- [ ] accessCondition
- [ ] part -> detail
- [ ] part -> extent
- [ ] part -> date
- [ ] part -> text
- [ ] extension
- [ ] recordInfo -> recordContentSource
- [ ] recordInfo -> recordCreationDate
- [ ] recordInfo -> recordChangeDate
- [ ] recordInfo -> recordIdentifier
- [ ] recordInfo -> recordOrigin
- [ ] recordInfo -> languageOfCataloging -> languageTerm
- [ ] recordInfo -> languageOfCataloging -> scriptTerm
- [ ] recordInfo -> descriptionStandard

Other (please specify)

If you use other elements not cited previously or other elements endorsed by other initiatives or created especially for your Repository, please specify by providing the metadata schema names and their elements.



Application Profiles and Metadata Schemes

CONCLUSION

Thank you for participating in this survey. Once the study is completed, I would be happy to share the results with you if you desire.

At what email address would you like to be contacted?

APÊNDICE B – Cartas enviadas para instituições e profissionais para participação no inquérito sobre TNis

Prezado(a) Senhor(a),

Sou estudante do terceiro ano do Doutoramento em Tecnologia e Sistemas de Informação da Universidade do Minho, Portugal, sob orientação da Professora Ana Alice Baptista, e estou a desenvolver uma investigação envolvendo a identificação do tipo de necessidade de informação dos utilizadores de serviços de informação. Os utilizadores selecionados para esta pesquisa são profissionais médicos e investigadores da área de **Oncologia**.

Gostaríamos de contar com a vossa colaboração na divulgação do estudo e do *link* para o questionário. Este *link* poderá ser disponibilizado por meio de redes sociais ou por *e-mail*. O questionário foi introduzido no *Survey*Monkey®, com tempo médio de resposta de 10 minutos, em que é preservado o anonimato. O *link* para o questionário é: https://pt.*survey*monkey.com/r/68JKV5M.

Para efeito de controlo do envio dos questionários, é interessante a identificação das instituições e o número de destinatários finais.

Agradeço o vosso apoio e a disponibilidade na contribuição para esta pesquisa. Caso tenha alguma dúvida, por favor, entre em contacto através do *e-mail*: morgana@dsi.uminho.pt ou whatsApp +351910514693 com Morgana Andrade.

Antecipadamente grata,

Morgana Andrade

Guimarães, 22 de abril de 2015

Prezado(a) Senhor(a),

Sou estudante do terceiro ano do Doutoramento em Tecnologia e Sistemas de Informação da Universidade do Minho, sob orientação da Professora Ana Alice Baptista, e estou a desenvolver uma pesquisa envolvendo a identificação do tipo de necessidade de informação dos utilizadores de serviços de informação. Os utilizadores selecionados para esta pesquisa são profissionais médicos, especialistas da área de Oncologia. Não foi identificado nenhum estudo do gênero em Portugal, diferente do que sucede em países como os Estados Unidos e a Grã-Bretanha.

Os médicos necessitam de respostas rápidas, precisas e confiáveis para responder às questões pontuais relacionadas a seus pacientes, já os investigadores, com o aumento do volume de informação e diversidade de sistemas de informação, também encontram dificuldades em obterem resultados com alto índice de relevância e precisão.

Este estudo visa propiciar aos médicos e investigadores, ao consultarem bases de dados bibliográficas, bibliotecas e repositórios digitais, o aumento da qualidade dos resultados da busca (maior relevância e precisão), com redução de esforços e do tempo utilizado pelos profissionais, aspectos extremamente relevantes para as atividades clínicas e de investigação, a partir do desenvolvimento de um artefacto que potencializa a troca de informação entre sistemas de informação, com base na semântica dos metadados.

Solicitamos a vossa colaboração na divulgação do estudo por meio de redes sociais ou media impressa (Newsletter), assim como de formas de contacto com os profissionais. O questionário foi introduzido no *Survey*Monkey (ferramenta online), cujo *link* poderá ser enviado por *e-mail*, disponibilizado no site do(a) ..... ou em rede social.

Agradeço o vosso apoio e a disponibilidade na contribuição para esta pesquisa. Caso tenha alguma dúvida, por favor, entre em contacto através do *e-mail*: morgana@dsi.uminho.pt ou telemóvel 910514693 com Morgana Andrade.

Antecipadamente grata,

Morgana Andrade



NECESSIDADES DE INFORMAÇÃO DOS ONCOLOGISTAS

Termo de Consentimento Livre e Esclarecido

Meu nome é Morgana Andrade, sou estudante do terceiro ano do Doutoramento em Tecnologia e Sistemas de Informação da Universidade do Minho. O presente estudo faz parte da elaboração da tese "Interoperabilidade semântica entre repositórios digitais científicos".
Neste âmbito estou a desenvolver um artefacto que potencializa a troca de informação entre sistemas de informação com base na semântica dos dados. Para isso é necessário identificar as necessidades de informação dos utilizadores de sistemas de informação.
Para este estudo, selecionamos especialistas da área de Oncologia (médicos e investigadores) que exercem a profissão no Brasil ou em Portugal.
A participação neste inquérito é voluntário. Os dados recolhidos para o presente estudo são de uso exclusivo do investigador e tratados de modo a garantir a sua máxima confidencialidade e anonimato.
Esta investigação tem como orientadora a Professora Doutora Ana Alice Baptista, responsável pela Disciplina Web Semântica e investigadora do Centro Algoritmi da Universidade do Minho.

* 1. Eu autorizo que os dados fornecidos sejam usados para este estudo e declaro que fui devidamente informado(a) e esclarecido(a).

- ◯ Sim
- ◯ Não



NECESSIDADES DE INFORMAÇÃO DOS ONCOLOGISTAS

DADOS PESSOAIS

* 2. Idade

- ◯ <35
- ◯ 35-44
- ◯ 45-54
- ◯ >55

* 3. Grau de formação mais alto

- ◯ Residência
- ◯ Especialização
- ◯ Mestrado
- ◯ Doutoramento
- ◯ Outro (especifique)

* 4. Qual a sua especialidade?

- ◯ Cirurgião oncológico
- ◯ Rádio-oncologista
- ◯ Oncologista clínico
- ◯ Outro (especifique)

* 5. Qual a sua atividade principal? Se necessário, indique mais de uma opção.

- [ ] Médico
- [ ] Pesquisador
- [ ] Professor
- [ ] Outro (especifique)

* 6. Quantos anos de experiência possui?

- ( ) < 5 anos
- ( ) 6 - 15 anos
- ( ) 16 - 25 anos
- ( ) 26 - 35 anos
- ( ) > 35 anos

7. Onde atua?

Cidade

Pais

Instituição (opcional)



NECESSIDADES DE INFORMAÇÃO DOS ONCOLOGISTAS

NECESSIDADE DE INFORMAÇÃO

Nesta secção procuraremos identificar as suas necessidades informacionais em relação à pesquisa de artigos científicos em serviços de informação (bibliotecas e repositórios digitais, bases de dados bibliográficas, etc). A necessidade de informação pode ser definida como um "estado ou processo iniciado quando alguém percebe que existe uma lacuna entre a informação e o conhecimento disponíveis para resolver um problema e a sua solução" (Miranda & Tarapanoff, 2008, p. 1).

8. Quando procura artigos quais dos itens discriminados abaixo utiliza?

- [ ] Título do artigo
- [ ] Título da publicação
- [ ] Autor
- [ ] Assunto
- [ ] Abreviaturas/siglas
- [ ] Figuras
- [ ] Títulos de figuras
- [ ] Tipo de estudo (revisão sistemática, estudo clínico randomizado)
- [ ] Suporte financeiro
- [ ] Factor de impacto
- [ ] Citações bibliográficas
- [ ] Tabelas
- [ ] Títulos de tabelas
- [ ] Notas (Uma nota ou anotação referente ao recurso)
- [ ] Análise de citações
- [ ] Utiliza ou gostaria de utilizar mais algum item?



NECESSIDADES DE INFORMAÇÃO DOS ONCOLOGISTAS

Tipos de necessidades de informação

As questões 8 a 16 podem estar relacionadas diretamente com a vossa atividade profissional: genética, diagnóstico, terapêutica, epidemiologia, patologia, imagiologia, assistência à saúde, pacientes. Informe apenas os termos relacionados com as categorias relevantes para si ao pesquisar artigos científicos em bases de dados bibliográficas, repositórios e bibliotecas digitais.

9. **Genética** (responda apenas se tiver necessidades de informação nesta categoria)

	Sempre	Às vezes	Raramente
Amount of target detected	○	○	○
Analysis technique and species	○	○	○
Anatomic location of target detected	○	○	○
Biological function (protein expression)	○	○	○
Biological process and species	○	○	○
Cell compartment and protein structure	○	○	○
Cell component	○	○	○
Chemical/organic compound	○	○	○
Comparison (tumor v adjacent normal)	○	○	○
Expression level (protein expression)	○	○	○
Gene expression experiments	○	○	○
Gene id (DNA)	○	○	○
Gene id (genetic information)	○	○	○
Gene/protein name and species and cell compartment	○	○	○
Methodology (affymetrix, northern blotting, tacman) (genetic information)	○	○	○
Methodology (genetic information)	○	○	○
Methodology (protein expression)	○	○	○
Name of protein (protein expression)	○	○	○
Plasmid	○	○	○
Probe type (genetic information)	○	○	○

	Sempre	Às vezes	Raramente
Processing e storage	○	○	○
Protein expression (genetic information)	○	○	○
Protein expression experiment (genetic information)	○	○	○
Protein id (genetic information)	○	○	○
Protein target (genetic information)	○	○	○
Sample annotation	○	○	○
Sample collection	○	○	○
Sample type (protein expression) (genetic information)	○	○	○
Sample-specific elements	○	○	○
Sequence (gene id (DNA))	○	○	○
Sequence (protein id)	○	○	○
Species and cell compartment	○	○	○
Species name and author	○	○	○
Species name and method	○	○	○
Structure (gene id (DNA)	○	○	○
Structure (protein id)	○	○	○
Tissue microarray (anatomic location of target detected (genetic information)	○	○	○
Tissue type and disease	○	○	○
Transcript (probe set) (genetic information)	○	○	○
Transcript id (RNA)	○	○	○

Sugere outros termos?

10. **Diagnóstico** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Anatomic region (specimen) (diagnostic procedure)	○	○	○
Body part (anatomy) on physical exam or imaging study	○	○	○
Cause/ interpretation of clinical finding	○	○	○
Cost	○	○	○
Criteria/ manifestations	○	○	○
Characteristics of a test	○	○	○
Date of procedure	○	○	○
Diagnosis, multiple	○	○	○
Diagnosis, using laboratory	○	○	○

	Sempre	Às vezes	Raramente
Diagnostic aid	○	○	○
Diagnostic criteria	○	○	○
Diagnostic procedures	○	○	○
Differential diagnosis	○	○	○
Disease accuracy (test)	○	○	○
How processed (specimen) (diagnostic procedure)	○	○	○
Image (specimen) (diagnostic procedures)	○	○	○
Images (diagnostic procedure)	○	○	○
Inconsistencies (diagnosis)	○	○	○
Indications/ efficacy (test)	○	○	○
Laboratory findings (images)	○	○	○
Laboratory and physiology findings	○	○	○
Method (test)	○	○	○
Morbidity/mortality of diagnostic procedure	○	○	○
Name finding (test)	○	○	○
Orientation (diagnosis)	○	○	○
Pathology findings (specimen)	○	○	○
Patient diagnosis	○	○	○
Physical signs interpretation	○	○	○
Preparation (test)	○	○	○
Sign (cause/interpretation of clinical finding)	○	○	○
Specimen (diagnostic procedures)	○	○	○
Symptom (cause/interpretation of clinical finding)	○	○	○
Test (lab, ECG, imaging, biopsy, skin test, element of physical exam, etc)	○	○	○
Test (orientation)	○	○	○
Test finding (lab, ECG, imaging, biopsy, skin test, etc.)	○	○	○
Timing/ monitoring (test)	○	○	○
Tumor grade, size and stage at DX and resection	○	○	○
Type of specimen (block, frozen, tissue, slides, serum) (diagnostic procedure)	○	○	○
Types of procedure (radiology, lab, surgery)	○	○	○
Unspecified findings or multiple categories of findings	○	○	○

Sugere outros termos?

11. **Terapêutica** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Administration in face of adverse effects	○	○	○
Adverse effects	○	○	○
Availability (treatment)	○	○	○
Cancer directed/non cancer directed (therapy)	○	○	○
Chemical/drug	○	○	○
Contraindications	○	○	○
Cost (drug prescribing)	○	○	○
Diet therapy	○	○	○
Dosage	○	○	○
Drug prescribing	○	○	○
Efficacy/ indications/ drug of choice	○	○	○
False-negative rate of diagnostic procedure	○	○	○
Findings caused by drug/ adverse effects of drug	○	○	○
How to do it (treatment)	○	○	○
Interactions (drug prescribing)	○	○	○
Mechanism of action (drug prescribing)	○	○	○
Modification,additions/omission to therapy	○	○	○
Name finding (drug prescribing)	○	○	○
Orientation/ composition (drug prescribing)	○	○	○
Outcomes and progression data – vital status and progression markers	○	○	○
Pharmaco-dynamics/ absorption (drug prescribing)	○	○	○
Physical characteristics (drug prescribing)	○	○	○
Prevention	○	○	○
Principles/ rationale (treatment)	○	○	○
Prior metastatic/recurrent therapy	○	○	○
Prior therapy	○	○	○
Protocol (therapeutic procedures)	○	○	○
Reactions, safety and	○	○	○
Reason for no-therapy	○	○	○
Reason for therapy ended	○	○	○
Safety/ contraindications (includes pregnancy and breast feeding)	○	○	○
Serum levels (drug prescribing)	○	○	○

	Sempre	Às vezes	Raramente
Start/end date (therapeutic procedures)	○	○	○
Therapy sequence	○	○	○
Timing (drug prescribing)	○	○	○
Treatment efficacy	○	○	○
Treatment recommendation	○	○	○
Treatment, adverse	○	○	○
Type of therapy (surgery, radiation)	○	○	○

Sugere outros termos?

12. **Epidemiologia** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Causation/ association (etiology)	○	○	○
Course/ prognosis	○	○	○
Etiology	○	○	○
Genetics	○	○	○
Population statistics	○	○	○
Prevalence/ incidence	○	○	○
Risk factors/ disease agents	○	○	○

Sugere outros termos?

13. **Patologia** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Cause and risk factors	○	○	○
Class of disease (cancer) (nature of disease)	○	○	○
Data of diagnosis (nature of disease)	○	○	○
Date of staging (extent of disease)	○	○	○
Diagnosis/etiology	○	○	○
Disease complication	○	○	○
Disease prognosis	○	○	○
Documentation for this cancer) (nature of disease)	○	○	○
Method of diagnosis (nature of disease)	○	○	○
Organ primary (class of disease)	○	○	○
Pathologic diagnosis (histological types)	○	○	○
Precursor lesions (nature of disease)	○	○	○
Specific findings TNM (stage) (extent of disease)	○	○	○
Stage (extent of disease)	○	○	○
Staging system (stage) (extent of disease)	○	○	○
Status at staging (extent of disease)	○	○	○
TNM stage (extent of disease)	○	○	○

Sugere outros termos?

14. **Imagiologia** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Anatomic frame of reference (image acquisition module)	○	○	○
Chemical agent administration (image acquisition module)	○	○	○
Color coding (images)	○	○	○
Color depth (images)	○	○	○
Compression (images)	○	○	○
Description of processing (image processing)	○	○	○
Equipment (images)	○	○	○
File location (physical description) (images)	○	○	○
File name (physical description) (images)	○	○	○
File size (physical description) (images)	○	○	○
Illumination (image acquisition module)	○	○	○

	Sempre	Às vezes	Raramente
Image acquisition context (image acquisition module)	◯	◯	◯
Image acquisition module	◯	◯	◯
Image center point coordinates sequence (slide coordinate module) (images)	◯	◯	◯
Image type (images) (light microscope, endoscopy, electron microscope)	◯	◯	◯
Lossy/lossless (images)	◯	◯	◯
Magnification (image acquisition module)	◯	◯	◯
Manufacturer of equipment (images)	◯	◯	◯
Monochrome2 or RGB or YBR_full_422 (images)	◯	◯	◯
Name of equipment (images)	◯	◯	◯
Physical description (images)	◯	◯	◯
Pixel resolution (the amount of specimen subtended by a pixel) (images)	◯	◯	◯
Primary magnification (images)	◯	◯	◯
Ratio (images)	◯	◯	◯
Related image sequence information (images)	◯	◯	◯
Size of pixel (images)	◯	◯	◯
Slide coordinate module (images)	◯	◯	◯
Software (equipment) (image)	◯	◯	◯
Software (image processing)	◯	◯	◯
Specimen acquisition and processing (image acquisition module)	◯	◯	◯

Sugere outros termos?

15. **Assistência à saúde** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Clinical trial information	○	○	○
Clinically condition (orientation)	○	○	○
Community services (management)	○	○	○
Condition/ finding (management)	○	○	○
Doctor-patient communication	○	○	○
How to advise (management)	○	○	○
How to approach difficult issue (doctor-patient communication)	○	○	○
Management (not specifying diagnostic or therapeutic) (management)	○	○	○
Patient compliance (doctor-patient communication)	○	○	○
Practices of other providers (management)	○	○	○
Specifying diagnostic or therapeutic) (management)	○	○	○

Sugere outros termos?

16. **Paciente (acompanhamento/resultados)** (responda apenas se tiver necessidades de informação nesta categoria):

	Sempre	Às vezes	Raramente
Autopsy	○	○	○
Cancer status at death	○	○	○
Date	○	○	○
Date last known alive	○	○	○
Date last verified recurrence–free	○	○	○
Date of death	○	○	○
Date of first recurrence	○	○	○
Date of last contact	○	○	○
Date of metastasis	○	○	○
Demographic data elements	○	○	○
Ethics	○	○	○
Family history of disease (family history of disease)	○	○	○
How documented (biopsy, cytology, radiology, physical exam)	○	○	○
How a patient should be managed disease information	○	○	○
Metastases (multiple)	○	○	○
Method used to determine first recurrence	○	○	○
Place of death	○	○	○
Primary cause of death	○	○	○
Quality of survival	○	○	○
Race and gender	○	○	○
Site of first distant recurrence	○	○	○
Site of local–regional recurrence	○	○	○
Site of metastasis	○	○	○
Type of first recurrence (local, distant, biochemical)	○	○	○
Vital status	○	○	○

Sugere outros termos?

17. Abaixo estão discriminadas categorias relacionadas com a sua atividade profissional, por favor indique nos respectivos campos as necessidades de informação que possui referente a estas categorias. Pode fazê-lo em forma de afirmações, perguntas ou tópicos, no idioma que preferir, p. ex.: What are the performance characteristics (sensitivity, specificity, etc.) of test "x" in situation "y"?.

Diagnóstico	
Terapêutica	
Procedimentos cirúrgicos	
Epidemiologia	
Compostos químicos e drogas	
Patologia	
Etiologia	
Imagiologia	
Técnicas de investigação (utilizada na pré-clínica, investigação clínica, epidemiologia, imunologia e genética)	
Equipamento e suplementos	
Sugere outras categorias?	



NECESSIDADES DE INFORMAÇÃO DOS ONCOLOGISTAS

CONCLUSÃO

18. Caso possua sugestões ou considerações em relação ao tema "necessidades de informação" ou "tipos de necessidades de informação", por favor indique no campo abaixo.

19. Gostaria de receber o resultado deste questionário, em caso afirmativo, preencha os dados que achar necessários.

Nome	
Instituição	
Endereço de email	
N° telefone	

20. Agradecemos o seu contributo que foi muito importante nesta fase da investigação. Na hipótese de haver necessidade efetuar entrevista, podemos contar com a sua participação? Em caso afirmativo, por favor indique seu email:



PROPOSTA DE UM MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Termo de Consentimento Livre e Esclarecido

Prezado(a) Senhor(a),

Convidamo-lo a participar de um Estudo Delphi, uma sondagem sistemática de opiniões de especialistas em determinada área, através de pesquisas interativas, na tentativa de chegar a um consenso de grupo sobre um determinado tópico. Este estudo considera a utilização de duas rondas.

O objetivo deste estudo Delphi é obter uma análise de classes e propriedades derivadas de uma investigação sobre tipos de necessidades de informação associados à busca por artigos científicos em repositórios digitais científicos e bases de dados bibliográficas no domínio da Oncologia.

Investigadora principal: Morgana Andrade, Universidade do Minho
Contato: morganaandrade@hotmail.com ou whatsapp +351910514693

Orientadora: Ana Alice Baptista, investigadora do Centro Algoritmi – Universidadedo Minho

* 1. Autorizo que os dados fornecidos sejam usados para este estudo e declaro que fui devidamente informado(a) e esclarecido(a).

- ◯ Sim
- ◯ Não

* 2. Identificação

Grau de formação	
Tempo de experiência	
Atividade principal (serviço de referência, catalogação, administração, etc.)	
Instituição onde atua	



PROPOSTA DE UM MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Identificação de propriedades

Apresentamos um modelo de domínio desenvolvido para representar as informações contidas em um artigo científico no domínio da Oncologia no contexto dos repositórios digitais científicos. Esclarecemos que o modelo de domínio "caracteriza os tipos de coisas descritas pelos metadados e seus relacionamentos" (Coyle & Baker, 2009, p. 1).

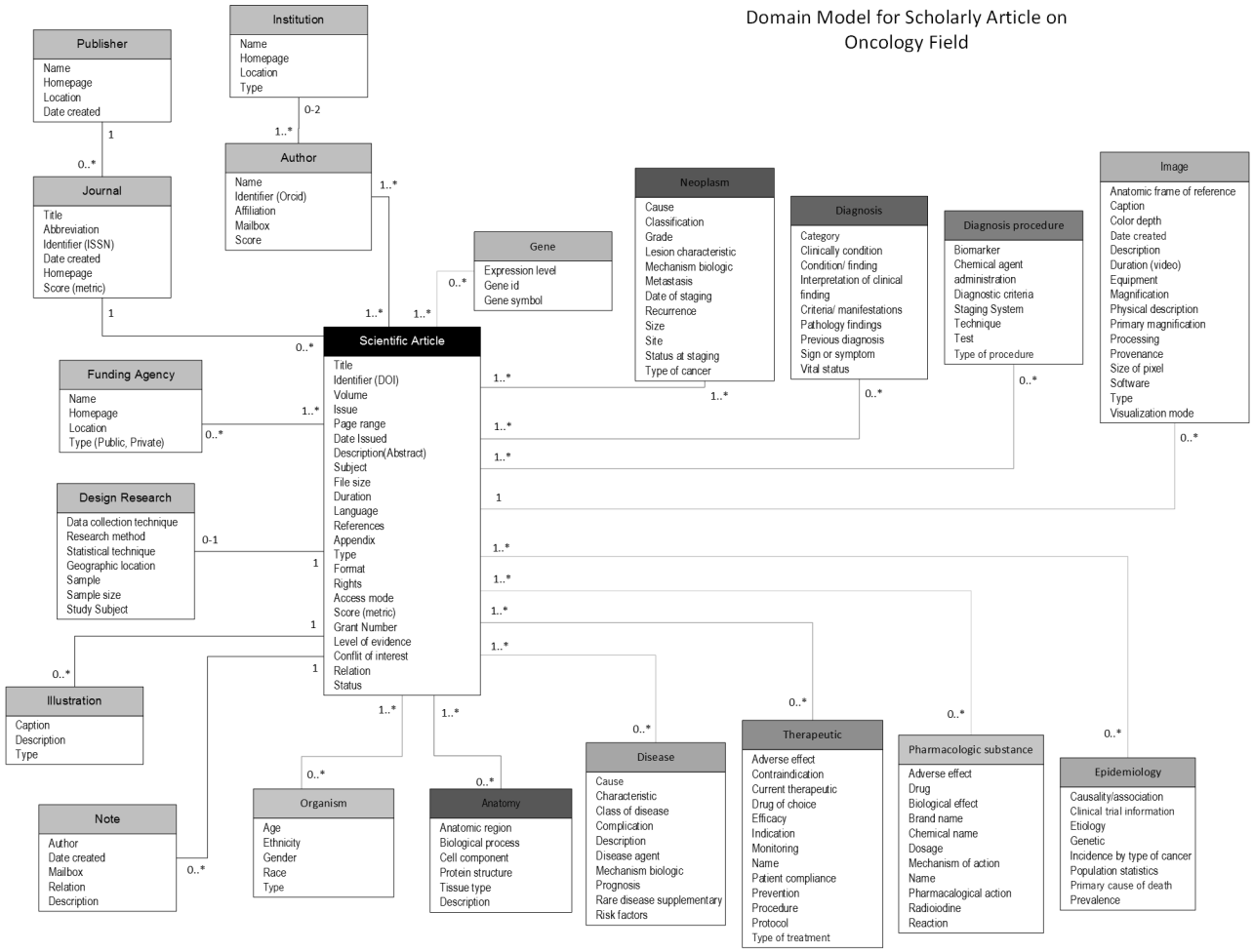
As informações foram representadas em um modelo de classes e propriedades. As classes correspondem a um conjunto de elementos que possuem propriedades, comportamentos, relacionamentos ou semântica em comum (POWELL et al., 2007) e as propriedades descrevem uma classe (DCMI, 2005). Exemplo: O artigo é uma classe; o título, uma propriedade desta classe.

Nesta secção, solicitamos que analise as classes e as propriedades associadas à pesquisa de artigos científicos em serviços de informação (bibliotecas e repositórios digitais, bases de dados bibliográficas etc). As caixas cinzas são classes que estão associadas à estrutura de um artigo científico e as caixas coloridas são classes que representam tópicos abordados por um artigo no domínio da Oncologia.

Gostaríamos de obter sua opinião sobre a representatividade dessas classes e suas propriedades para a descrição de artigos científicos avaliando:

a) se as propriedades estão devidamente representados nas classes ou se alguma delas estariam mais bem representadas em outra classe (Quadro 1); e/ou

b) casos de ambiguidade (pode ter mais do que um sentido ou significado) ou redundância (termos distintos para propriedades, porém com o mesmo significado e, portanto, o mesmo conteúdo). Nestes casos é solicitado que a propriedade deva ser eliminada.



* 3. **Classe:** Publisher (Publishing company)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name (A human-readable name for the subject)	○	○	○	○
Homepage (A homepage for publisher)	○	○	○	○
Location (The spatial topic of the resource)	○	○	○	○
Date created (Date of formal issuance of the resource)	○	○	○	○

Deixe seu comentário:

* 4. **Classe:** Journal (A periodical of scholarly journal)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Title (A name given to the journal. Ex.: Advances in cancer research)	○	○	○	○
Abbreviation (The abbreviated form representing a resource. Ex.: Adv. cancer res.)	○	○	○	○
Identifier (An unambiguous reference to the resource. Ex.: ISSN: 0065-230X)	○	○	○	○
Publisher (An entity responsible for making the resource available. Ex.: https://www.elsevier.com.br/)	○	○	○	○
Date Created (Date of formal issuance of the resource. Ex.: 1953-03-23)	○	○	○	○
Homepage (A homepage for resource. Ex.: http://www.sciencedirect.com/science/bookseries/0065230X).	○	○	○	○
Score (Represents the metric for resource. Ex.: 82 H Index (http://www.scimagojr.com/journalsearch. php?q =28690&tip=sid)	○	○	○	○

Deixe seu comentário:

* 5. **Classe:** Scientific Article (A scholarly academic article, typically published in a journal)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a essa classe	Não sei
Title (A name given to the resource)	○	○	○	○
Identifier (An unambiguous reference to the resource within a given context, ex. DOI , Handle)	○	○	○	○
Volume (Volume of a journal)	○	○	○	○
Issue (Issue of a journal)	○	○	○	○
Page range (Page range of a document)	○	○	○	○
Date Issued (Date of publication of the resource)	○	○	○	○
Author (An entity primarily responsible for making the resource. Ex.: 000-0002-4055-5882)	○	○	○	○
Abstract (A summary of the resource).	○	○	○	○
Subject (The topic of the resource).	○	○	○	○
Language (A language of the resource)	○	○	○	○
References (A related resource that is referenced, cited by the described resource)	○	○	○	○
Appendix (Text or document prepared by the author)	○	○	○	○
Illustration (A numbered and captioned image, graphic or table within a journal)	○	○	○	○

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a essa classe	Não sei
Type (The nature or genre of the resource. Ex.: Research article, case report, etc.)	○	○	○	○
Format (The file format, physical medium, or dimensions of the resource. Ex.: html, xml, pdf, doc, mp3, jpeg, mpeg)	○	○	○	○
File size corresponding to the file of article (ex.: 300MB)	○	○	○	○
Duration of the audio or video corresponding to the file of article (ex.: 30 min)	○	○	○	○
Rights (Information about rights held in and over the resource. Ex.: Licence Creative Commons)	○	○	○	○
Access Mode (Declares the degree of openness of a resource. Ex.: open, embargo or restrict)	○	○	○	○
Score (Represents the metric being observed. Ex.: Scopus, Scholar Google, Web of Science)	○	○	○	○
Funding Agency (Institutions that offer different forms of financing aimed at conducting research. Ex.: Fundação para a Ciência e a Tecnologia (FTC - Portugal)	○	○	○	○
GrantNumber (A data property specifying the grant number of a grant provided by a funding agency. Ex.: UID/Multi/04378/2013)	○	○	○	○
Level of Evidence (Strength of evidence of the data used to formulate the guideline. Ex.: Level 1A for treatment corresponding a systematic review of randomized clinical trial)	○	○	○	○
Conflit of Interest (A situation in which an individual might benefit personally from oficial or professional actions)	○	○	○	○
Relation (A related resource. Ex.: a thesis, a research project)	○	○	○	○
Status (Status of a resource. Ex.: pre print; post print)	○	○	○	○

Deixe seu comentário:

* 6. **Classe:** Author (The individual who is responsible for the content of a document)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name (An entity primarily responsible for making the resource)	◯	◯	◯	◯
Identifier (An unambiguous reference to the resource within a given context. Ex.: Orcid)	◯	◯	◯	◯
Affiliation (An organization that this person is affiliated)	◯	◯	◯	◯
Mailbox (A personal mailbox)	◯	◯	◯	◯
Score (A numeric values measuring performance or quality of something. Ex.: Google Citation)	◯	◯	◯	◯

Deixe seu comentário:

* 7. **Classe:** Institution (An organization, establishment or foundation)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name (Name of Institution. Ex.: Universidade do Minho)	◯	◯	◯	◯
Homepage (A homepage for organization)	◯	◯	◯	◯
Location (Physical address of the item)	◯	◯	◯	◯
Type (The nature or genre of the resource. Ex.: Academic institution, hospital, etc.)	◯	◯	◯	◯

* 8. **Classe:** Funding Agency (Instiutions or organizations that support scientific or technologic research)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Name (Name of agency. Ex.: FCT)	◯	◯	◯	◯
Homepage (A homepage for organization)	◯	◯	◯	◯
Location (Physical address of the item)	◯	◯	◯	◯
Type (The nature or genre of the resource. Ex.: public, private)	◯	◯	◯	◯

Deixe seu comentário:

* 9. **Classe:** Illustration (A numbered and captioned image, graphic or table within a journal)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Caption (The caption for this object)	◯	◯	◯	◯
Description (An account of the resource)	◯	◯	◯	◯
Type (The nature or genre of the resource. Ex.: graphic, table or figure)	◯	◯	◯	◯

Deixe seu comentário:

* 10. **Classe:** Note (A note about a resource)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Author (Na entity primarily responsible for making the resource. Ex.: Smith, J.)	◯	◯	◯	◯
Date created (Date of creation of the resource)	◯	◯	◯	◯
Mailbox (A personal mailbox)	◯	◯	◯	◯
Relation (A related resource)	◯	◯	◯	◯
Description (An account of the resource)	◯	◯	◯	◯

* 11. **Classe:** Design Research (Describes how the study performed to answer the research)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Data collection technique (Referes to process of gathering data for testing or research purposes)	◯	◯	◯	◯
Research method (A method or procedures used in a research)	◯	◯	◯	◯
Statistical technique (A statistical technique used in a research)	◯	◯	◯	◯
Geographic location (Place where the study was developed)	◯	◯	◯	◯
Sample (The sample of study)	◯	◯	◯	◯
Sample size (The sample size)	◯	◯	◯	◯
Study Subject (A subject investigated by the study. Ex.: Therapy, Drug, Diagnosis)	◯	◯	◯	◯

Deixe seu comentário:

* 12. **Classe:** Gene (The functional and physical unit of heredity passed from parente to offspring)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Expression level (The determination of the level of gene transcription in a biological specimen)	○	○	○	○
Gene id (A unique gene name within a specific repository, database, or collection)	○	○	○	○
Gene symbol (A unique gene name approved by an organism specific nomenclature committee)	○	○	○	○

Deixe seu comentário:

* 13. **Classe:** Organism (A living entity)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Age (Elapsed time since birth)	○	○	○	○
Ethnicity (Members share a common history and a sense of identification with the group)	○	○	○	○
Gender (The gender of this agent. Ex.: male or female)	○	○	○	○
Race (An arbitrary classification of a taxonomic group that is a division of a species)	○	○	○	○
Description (Description of a living entity. Ex.: Animal or human being)	○	○	○	○

Deixe seu comentário:

* 14. **Classe:** Anatomy (Physical structure of animals)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Anatomic region (Identifies areas of the body)	◯	◯	◯	◯
Biological process (An activity occurring within an organism, between organisms or among organisms and the mechanisms underlying such events)	◯	◯	◯	◯
Cell component (The components of the cell)	◯	◯	◯	◯
Protein structure (The amino acids and their 3-D arrangement in constituting a protein)	◯	◯	◯	◯
Tissue type (Describes collections of differentiated cells, such as epithelium; connective tissue; muscles; nerve tissue)	◯	◯	◯	◯
Description (Describes the referenced anatomy)	◯	◯	◯	◯

Deixe seu comentário:

* 15. **Classe:** Neoplasm (A benigno or malignant tissue growth resulting from uncontrolled cell proliferation)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Cause (One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence)	○	○	○	○
Classification (The classification of cancers by histological type; primary site; location where it was first developed. Ex.: 8890/0 LEIOMYOMA, NOS or http://codes.iarc.fr/code/3406)	○	○	○	○
Date of staging (Date of he process of defining at what point in the natural history of a malignant disease a patient is when a diagnosis is made)	○	○	○	○
Grade (A position on a scale of intensity or amount or quality. Ex.: 8077/2 Squamous intraepithelial neoplasia, High Grade)	○	○	○	○
Lesion characteristic (Characteristic of localized pathological or traumatic structural change, damage, deformity, or discontinuity of tissue, organ, or body part)	○	○	○	○
Mechanism biologic (Mechanism biologic associated with cancer)	○	○	○	○
Metastasis (The spread or migration of cancer cells from one part of the body to another)	○	○	○	○
Recurrence (If the cancer returned)	○	○	○	○
Site (A position in which cancer develops)	○	○	○	○
Size (The physical magnitude of tumor)	○	○	○	○
Status at staging (Level of extent of the cancer within the body)	○	○	○	○
Type of cancer (Type of cancer. Ex.: Colorectal cancer, endometrial cancer)	○	○	○	○

Deixe seu comentário:

* 16. **Classe:** Disease (A health problem with certain characteristics or symptoms)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Cause (One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence)	○	○	○	○
Characteristic (The distinguishing of a disease aspect)	○	○	○	○
Class of disease (Classification of diseases. Ex.: Breast cancer)	○	○	○	○
Complication (Any disease or disorder that occurs during the course of, or because of, another disease, treatment, or procedure)	○	○	○	○
Description (Description of the disease)	○	○	○	○
Disease agent (Any vírus, micro-organism, or other substance causing disease)	○	○	○	○
Mechanism biologic (Mechanism biologic associated with disease)	○	○	○	○
Prognosis (Prognosis)	○	○	○	○
Rare disease supplementary (Rare disease identification).	○	○	○	○
Risk factors (A modifiable or non-modifiable fator that increases the risk of a person contracting a condition. Ex.: Tobacco)	○	○	○	○

Deixe seu comentário:

* 17. **Classe:** Diagnosis (The investigation, analysis and recognition of the presence and nature of disease, condition, or injury from expressed signs and symptoms)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Category (A category for the item. Ex.: Category based on image clinical report – BI-RADS)	○	○	○	○
Clinically condition (Specifying the health condition(s) of a patient)	○	○	○	○
Condition/ finding (Clinical, laboratory or molecular evidence, or lack of evidence of the disease)	○	○	○	○
Interpretation of clinical finding (Interpretation of clinical finding)	○	○	○	○
Criteria/ manifestations (Criteria based on patient symptoms and signs, test results. Ex.: Amsterdam criteria)	○	○	○	○
Pathology findings (Pathology findings)	○	○	○	○
Previous diagnosis (Previous diagnosis)	○	○	○	○
Sign or symptom (A sign o r symptom of this condition)	○	○	○	○
Vital status (The state or condition of being living or deceased; also includes the case where the vital status is unknown)	○	○	○	○

Deixe seu comentário:

* 18. **Classe:** Diagnostic procedure (A medical procedure intented primarily for diagnostic, as opposed to therapeutic, purposes)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Biomarker (Biomarker used)	○	○	○	○
Chemical agent administration (Administration of chemical agent in exams)	○	○	○	○
Diagnostic criteria (A standard from which a judgment concerning a medical diagnosis can be established. Ex.: Amsterdam criteria)	○	○	○	○
Staging system (A system adoped to describe the extent of cancer in a patient's body. Ex.: TNM Staging System)	○	○	○	○
Technique (Encompass all investigations and tests intended to identify the cause of disease. Ex.: laboratory tests, Differential diagnosis; Imaging techniques)	○	○	○	○
Test (Process of determining the disease state of a person. Ex.: cancer mutation test, blood chemistry test, etc.)	○	○	○	○
Type of procedure (The type of procedure, for example surgical, noninvasive or percutaneous. Ex.: Clinical diagnosis; computer-aided diagnosis).	○	○	○	○

Deixe seu comentário:

* 19. **Classe:** Image (An image)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Anatomic frame of reference (Image acquisition module)	○	○	○	○
Caption (The caption for this object)	○	○	○	○
Color depth (Positive integer indicating the desired minimum total number of bits per color channel used to present a pixel)	○	○	○	○
Date created (Date of creation of the resource)	○	○	○	○
Description (An account of the image)	○	○	○	○
Equipment (Equipment used to analyze, enhance, interpret or display digital images)	○	○	○	○
Format (The duration of the item. Ex.: For a video, T19:20+01:00)	○	○	○	○
Magnification (Magnification is related to scaling up visuals or images to be able to see more detail)	○	○	○	○
Physical description (Physical description of the image)	○	○	○	○
Primary magnification (Initial magnification)	○	○	○	○
Processing (The use of data processing systems to analyze, enhance, interpret or display digital images)	○	○	○	○
Provenance (Ownership and custody of the resource of since its creation. Ex.: Hospital, Laboratory)	○	○	○	○
Size of pixel (Image size expressed in pixels)	○	○	○	○
Software (A computer program in source of compiled form)	○	○	○	○
Type (The nature or genre of the resource. Ex.: radiography, videolaryngoscopy)	○	○	○	○
Visualization Mode (Options for visualization. Eex.: 3D image).	○	○	○	○

Deixe seu comentário:

* 20. **Classe:** Epidemiology (The study of the patterns, causes, and control of disease in groups of people)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Causality/ association (Relationship exists between a cause or effect. Ex.: radiation)	○	○	○	○
Clinical trial information (Report of a pre-planned medical or veterinary clinical case study involving a randomized, evidence-controlled clinical trial)	○	○	○	○
Etiology (The cause of a disease or abnormal condition)	○	○	○	○
Genetic (Interdisciplinary clinical, epidemiologic, genetic, and laboratory studies of persons, families, and populations at high risk of cancer)	○	○	○	○
Incidence by type of cancer (The relative frequency of occurrence of cancer)	○	○	○	○
Population statistics (A value, which represents a particular population characteristic)	○	○	○	○
Prevalence (The total number of cases of a given disease in a specified population at a designated time)	○	○	○	○
Primary cause of death (The first significant event which ultimately led to death)	○	○	○	○

Deixe seu comentário:

* 21. **Classe:** Therapeutic (Relating to or involved in therapy)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Adverse effect (An unexpected medical problem that happens during treatment with a drug or other therapy)	◯	◯	◯	◯
Contraindication (A contraindication for this therapy)	◯	◯	◯	◯
Current therapeutic (The name of the therapy that is currently being administered)	◯	◯	◯	◯
Drug of choice (The class of drug this belongs to. Ex., statins)	◯	◯	◯	◯
Efficacy (The ability of a chemical agent or a device to treat a disease)	◯	◯	◯	◯
Indication (A fator that indicates use of this therapy for treatment and/or prevention of a condition, symptom, etc.)	◯	◯	◯	◯
Monitoring (How a person or a condition is assited or verified)	◯	◯	◯	◯
Name (A human-readable name for the subject)	◯	◯	◯	◯
Patient compliance (The act of following a medical regimen or schedule correctly and consistently, including taking medicines or following a diet)	◯	◯	◯	◯
Prevention (Describes methods of preventing cancer including limit of exposure to carcinogens and infectious agents, diet, education, etc)	◯	◯	◯	◯
Procedure (Describes a treatment or action taken to prevent or treat disease)	◯	◯	◯	◯
Protocol (Describes a rule which guides how an activity should be performed)	◯	◯	◯	◯
Type of treatment (Cancer treatment adopted to non directed or directed cancer Ex.: control, remove or destroy cancer or prolong a patient’s live, alleviate pain)	◯	◯	◯	◯

Deixe seu comentário:

* 22. **Classe:** Pharmacologic Substance (Any substance, other than food, that is used to prevent, diagnose, treat or relieve symptoms of a disease or abnormal condition)

	Deve ser mantida como propriedade	Não deve ser mantida como propriedade	Não está associada a esta classe	Não sei
Adverse effects (An unexpected medical problem that happens during treatment with a drug or other therapy)	○	○	○	○
Drug (Specifying a drug or medicine used in a medication procedure)	○	○	○	○
Biological effects (Biological effect that the any product provide. Ex.: Radiation)	○	○	○	○
Brand name (Brand name. Ex.: Ibuprofen 400 to Ibuprofen)	○	○	○	○
Chemical name (Chemical name. Ex.: Ibuprofen; Brufen to Ibuprofen)	○	○	○	○
Dosage (The value of the dose. Ex. 500 mg)	○	○	○	○
Mechanism of action (The specific biochemical interaction through which this drug its pharmacological effect)	○	○	○	○
Pharmacalogical action (Actions of drugs and their effects on living organisms)	○	○	○	○
Radioiodine (Use of radioiodine)	○	○	○	○
Reaction (A description of the action in response to some kind of stimulus)	○	○	○	○

Deixe seu comentário:



PROPOSTA DE UM MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Conclusão

23. Considerando este modelo de domínio, indique classes que você considera que não foram representadas.

24. Você considera que este modelo de domínio atende aos requisitos de busca para um artigo no domínio da Oncologia?

- ◯ Sim
- ◯ Não
- ◯ Justifique:

* 25. O preenchimento do endereço de e-mail ou número do whatsapp é necessário para a segunda ronda do Estudo Delphi. Agradecemos o seu contributo que foi muito importante nesta fase da investigação.

Endereço de email

Whatsapp



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Estudo Delphi - Segunda Ronda

Prezado(a) Bibliotecário(a),

A primeira ronda gerou resultados que foram sintetizados e precisam ser reavaliados por todos os respondentes, a fim de obter um consenso. Para melhor entender o processo, peço que dedique alguns instantes à leitura do texto a seguir.

Na avaliação anterior, foi possível identificar o percentual de respostas em relação a: manutenção das propriedades em determinada classe; não manutenção da propriedade; a propriedade estar relacionada a outra classe.

O presente questionário foi elaborado de modo que os participantes tenham conhecimento sobre esse resultado, nomeadamente em relação às propriedades que apresentaram alguma divergência na avaliação. No sentido de dirimir alguma dúvida, serão adicionados esclarecimentos suplementares. Lembrando que o modelo em discussão descreve um artigo científico no contexto dos repositórios digitais científicos.

Este questionário foi estruturado em 18 páginas e cada página corresponde a uma classe. São apresentadas as propriedades, já avaliadas por você e pelos outros respondentes, e os percentuais obtidos.

A cada questão, serão fornecidas faixas de percentuais quanto a manutenção da propriedade: a) 100% a favor da manutenção da propriedade, logo não há necessidade de manifestação por parte do grupo; e b) 88%, 78% e igual ou inferior a 63%. Diante desses índices, será solicitado aos respondentes que as propriedades sejam reavaliadas ("concordo fortemente", “concordo” e “não concordo”).

Informamos, ainda, que as propriedades serão mantidas ao final da segunda ronda se houver um índice de aprovação igual ou superior a 75%, índice preconizado por Murry & Hammons (1995).

Agradecemos, mais uma vez, por sua contribuição para a conclusão deste estudo.

Investigadora principal: Morgana Andrade, Universidade do Minho
Contato: morganaandrade@hotmail.com ou whatsapp +351910514693

Orientadora: Ana Alice Baptista, investigadora do Centro Algoritmi – Universidadedo Minho



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Publisher (Publishing company)

A propriedade a seguir obteve um percentual de 100% de aprovação:

- **Name**

* 1. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Homepage - representa a homepage do Journal	◯	◯	◯
Location - representa a identificação da localização da editora	◯	◯	◯
Date created - representa a data de criação do Journal.	◯	◯	◯

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Journal (A periodical of scholarly journal)

As propriedades abaixo obtiveram um percentual de 100% de aprovação:

- **Title**
- **Identifier**
- **Publisher**
- **Date Created**
- **Homepage**

Score

* 2. A propriedade a seguir obteve 88% de aprovação. Você concorda que ela deve ser mantida?

	Concordo fortemente	Concordo	Não concordo
Abbreviation (representa o título abreviado do Journal, utilizado por muitos estilos de referências bibliográficas)	○	○	○

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Scientific Article (A scholarly academic article, typically published in a journal)

As propriedades abaixo obtiveram um percentual de 100% de aprovação:

- **Title**
- **Identifier**
- **Volume**
- **Issue**
- **Date Issued**
- **Author**
- **Abstract**
- **Level of Evidence**

* 3. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Page range (refere-se ao intervalo de páginas do artigo, inicial e final)	◯	◯	◯
Subject (representa as palavras-chave utilizadas para descrever o artigo)	◯	◯	◯
Language (representa o idioma adotado pelo artigo)	◯	◯	◯
References (representa as referências utilizadas no artigo)	◯	◯	◯
Format (representa o formato do arquivo, ex.: pdf, word)	◯	◯	◯
Access Mode (representa ao tipo de acesso ao artigo, ex.: acesso aberto, embargado ou restrito)	◯	◯	◯
Score (representa o número de citações que o artigo obteve)	◯	◯	◯

Comente:

* 4. As propriedades a seguir obtiveram 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Rights (refere-se aos direitos autorais, nomeadamente à Licença Creative Commons)	◯	◯	◯
Conflit of Interest (refere-se a possíveis conflitos de interesses informados pelos autores)	◯	◯	◯
Relation (representa o(s) trabalho(s) que pode(m) ter originado o artigo, como teses, projetos de pesquisa)	◯	◯	◯
Status (representa o estágio do artigo ao ser armazenado no repositório, preprint, posprint)	◯	◯	◯
Illustration (refere-se aos títulos das ilustrações)	◯	◯	◯
Type (representa o tipo de artigo de revisão, pesquisa, ou um relato de caso)	◯	◯	◯

Comente:

* 5. As propriedades a seguir obtiveram um índice igual ou inferior a 63% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Appendix (representa os apêndices do artigo, que podem ser imagens, planilhas de dados)	◯	◯	◯
Funding Agency (representa agência de fomento que contribuiu para o desenvolvimento do estudo descrito pelo artigo)	◯	◯	◯
GrantNumber (refere-se ao número do processo do auxílio financeiro para o desenvolvimento do estudo)	◯	◯	◯
Duration (representa a duração do audio/vídeo para quando o artigo é disponibilizado em um desses formatos)	◯	◯	◯
File size (refere-se ao tamanho do arquivo do artigo)	◯	◯	◯

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Author (The individual who is responsible for the content of a document)

As propriedades a seguir obtiveram um percentual de 100% de aprovação:

- **Name**
- **Mailbox**

* 6. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Identifier (refere-se a identificação única do autor, o ORCID. Vem sendo utilizado por muitas editoras e recentemente como requisito para atribuição de financiamentos de pesquisa)	◯	◯	◯
Affiliation (refere-se ao vínculo empregatício do(s) autor(es)	◯	◯	◯
Score (refere-se a quantidade de citações do autor(es))	◯	◯	◯

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Institution (An organization, establishment or foundation)

A propriedade abaixo obteve um percentual de 100% de aprovação:

- **Name**

* 7. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Homepage (representa a homepage da instituição que o autor tem vínculo)	◯	◯	◯
Location (refere-se a localização da instituição a qual o autor está vinculado)	◯	◯	◯
Type (representa o tipo de instituição, ex.: hospital, universidade, instituto de pesquisa)	◯	◯	◯

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Funding Agency (Instiutions or organizations that support scientific or technologic)

A propriedade abaixo obteve 100% de aprovação:

- **Name**

* 8. As propriedades a seguir obtiveram 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Homepage (representa a homepage da agência de fomento)	◯	◯	◯
Location (refere-se a localização da agência)	◯	◯	◯
Type (refere-se ao tipo de organização, ex.: governamental, privada, etc.)	◯	◯	◯

Comente



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Design Research (Describes how the study performed to answer the research)

As propriedades abaixo obtiveram um percentual de 100% de aprovação:

- **Research method**
- **Statistical technique**
- **Study Subject**

* 9. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Data collection technique (refere-se as técnicas de coleta de dados descritas no artigo, ex.: questionário, entrevista, etc.)	◯	◯	◯
Geographic location (refere-se ao(s) local(is) onde foi desenvolvido o estudo descrito pelo artigo)	◯	◯	◯

Comente:

* 10. As propriedades a seguir obtiveram 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Sample (refere-se a amostra descrita no estudo)	◯	◯	◯
Sample size (refere-se ao tamanho da amostra descrita pelo estudo)	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Illustration (A numbered and captioned image, graphic or table)

Uma ilustração corresponde a figuras, quadros, etc.

* 11. As propriedades a seguir obtiveram 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Caption (refere-se aos títulos das ilustrações)	○	○	○
Type (representa o tipo de ilustração, ex.: figura, quadro, etc.)	○	○	○

Deixe seu comentário:

* 12. A propriedade a seguir obteve 63% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Description (refere-se a descrição do conteúdo da ilustração)	○	○	○

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Note (A note about a resource)

A classe Nota representa uma nota ou um comentário que os usuários podem atribuir ao artigo

* 13. A propriedade a seguir obteve um índice de 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Relation (inclui o link para o recurso)	◯	◯	◯

Outro (especifique)

* 14. As propriedades a seguir obtiveram 63% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Author (representa o autor do comentário ao artigo)	◯	◯	◯
Date created (representa a data em que o comentário foi feito)	◯	◯	◯
Mailbox (representa o e-mail, quando disponibilizado, pelo autor do comentário)	◯	◯	◯
Description (refere-se ao conteúdo do comentário)	◯	◯	◯

Outro (especifique)



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Gene (The functional and physical unit of heredity passed from parente to offs)

* 15. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Expression level (representa o nível de transcrição genética em uma amostra biológica)	◯	◯	◯
Gene id (refere-se à identificação dos genes descritos no artigo. Trata-se de uma identificação única dentro de um repositório, base de dados ou coleção)	◯	◯	◯
Gene symbol (refere-se ao nome do(s) gene(s) descrito(s) no artigo. Trata-se de um nome aprovado por um comitê específico)	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Organism (A living entity)

As propriedades a seguir obtiveram um percentual de 100% de aprovação:

- **Ethnicity**
- **Gender**
- **Description**

* 16. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Age (refere-se à idade dos participantes da investigação descrita pelo artigo)	◯	◯	◯
Race (refere-se à raça dos participantes da investigação descrita pelo artigo)	◯	◯	◯

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Anatomy (Physical structure of animals)

As propriedades abaixo obtiveram um percentual de 100% de aprovação:

- **Anatomic region**
- **Tissue type**

* 17. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Biological process (represents an activity occurring within an organism, between organisms or among organisms and the mechanisms underlying such events)	◯	◯	◯
Cell component (represents the components of the cell)	◯	◯	◯
Protein structure (represents the amino acids and their 3-D arrangement in constituting a protein)	◯	◯	◯
Description (represents a referenced anatomy)	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Neoplasm (A benigno or malignant tissue growth resulting from uncontrolled

As propriedades abaixo obtiveram um percentual de 100% de aprovação:

- **Classification**
- **Lesion characteristic**
- **Mechanism biologic**
- **Metastasis**
- **Recurrence**
- **Status at staging**
- **Type of cancer**

* 18. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Cause (represents one of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence) - Deve ser mantida	◯	◯	◯
Date of staging (represents date of he process of defining at what point in the natural history of a malignant disease a patient is when a diagnosis is made) - Deve ser mantida	◯	◯	◯
Grade (represents a position on a scale of intensity or amount or quality)	◯	◯	◯
Site (represents a position in which cancer develops in relation to morphology)	◯	◯	◯
Size (represents the physical magnitude of tumor)	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Disease (A health problem with certain characteristics or symptoms)

As propriedades a seguir obtiveram um percentual de 100% de aprovação:

- **Cause**
- **Characteristic**
- **Class of disease**
- **Complication**
- **Disease agent**
- **Mechanism biologic**
- **Prognosis**
- **Risk factors**

* 19. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Description (represents a description of the disease)	◯	◯	◯
Rare disease supplementary (represents a rare disease identification) -	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Epidemiology (The study of the patterns, causes, and control of disease in groups of people

As propriedades a seguir obtiveram um percentual de 100% de aprovação:

- **Causality/ association**
- **Etiology**
- **Genetic**
- **Incidence by type of cancer**
- **Population statistics**
- **Prevalence**
- **Primary cause of death**

* 20. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Clinical trial information (represents a report of a pre-planned medical or veterinary clinical case study involving a randomized, evidence-controlled clinical trial)	○	○	○

Comente:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Diagnosis (The investigation, analysis and recognition of the presence and nature

A propriedade a seguir obteve um percentual de 100% de aprovação:

- **Pathology findings**

* 21. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Category (represents a category for the item)	○	○	○
Clinically condition (represents a specifying the health condition(s) of a patient)	○	○	○
Condition/ finding (represents a clinical, laboratory or molecular evidence, or lack of evidence of the disease)	○	○	○
Interpretation of clinical finding (represents an interpretation of clinical finding)	○	○	○
Criteria/ manifestations (represents a criteria based on patient symptoms and signs, test results)	○	○	○
Previous diagnosis	○	○	○
Vital status (represents a state or condition of being living or deceased; also includes the case where the vital status is unknown)	○	○	○

Deixe seu comentário:

* 22. As propriedades a seguir obtiveram 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Sign or symptom (refere-se aos sinais e sintomas descritos no artigo quanto a determinada doença)	○	○	○

Comete:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Diagnostic procedure (A medical procedure intented primarily for diagnostic, as opposed to therapy)

As propriedades abaixo obtiveram um percentual de 100% de aprovação:

- **Chemical agent administration**
- **Test**
- **Type of procedure**

* 23. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Biomarker (represents biomarker used)	○	○	○
Diagnostic criteria (represents a standard from which a judgment concerning a medical diagnosis can be established)	○	○	○
Staging system (represents a system adoped to describe the extent of cancer in a patient's body)	○	○	○
Technique (represents all investigations and tests intended to identify the cause of disease)	○	○	○

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Image (An image)

A classe Image representa um elemento gráfico descrito por um artigo científico

* 24. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Anatomic frame of reference (represents an image acquisition module)	◯	◯	◯
Caption (represents the caption for this object)	◯	◯	◯
Date created (represents a date of creation of the resource)	◯	◯	◯
Description (represents an account of the image)-	◯	◯	◯
Magnification (represents a magnification is related to scaling up visuals or images to be able to see more detail)	◯	◯	◯
Provenance (represents a ownership and custody of the resource of since its creation)	◯	◯	◯
Type (represents the nature or genre of the resource)	◯	◯	◯
Visualization Mode (represents an options for visualization)	◯	◯	◯

Deixe seu comentário:

* 25. As propriedades a seguir obtiveram 75% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Equipment (represents an equipment used to analyze, enhance, interpret or display digital images)	◯	◯	◯
Format (represents the duration of the item)	◯	◯	◯
Physical description (represents a physical description of the image)	◯	◯	◯
Processing (represents the use of data processing systems to analyze, enhance, interpret or display digital images)	◯	◯	◯
Software (represents a computer program in source of compiled form)	◯	◯	◯

Deixe seu comentário:

* 26. As propriedades a seguir obtiveram índice igual ou inferior a 69% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Color depth (represents a positive integer indicating the desired minimum total number of bits per color channel used to present a pixel)	◯	◯	◯
Primary magnification (represents an initial magnification)	◯	◯	◯
Size of pixel (represents an image size expressed in pixels)	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Therapeutic

As propriedades a seguir obtiveram um percentual de 100% de aprovação:

- **Adverse effect**
- **Contraindication**
- **Current therapeutic**
- **Drug of choice**
- **Efficacy**
- **Indication**
- **Monitoring**
- **Patient compliance**
- **Procedure**
- **Protocol**
- **Type of treatment**

* 27. A propriedade a seguir obteve 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Prevention (represents the methods of preventing cancer including limit of exposure to carcinogens and infectious agents, diet, education, etc)	○	○	○

Comente

* 28. Foi sugerida a inclusão da classe Psychological Support, após avaliação, optou-se por incluí-la como propriedade. Você concorda com essa inclusão?

	Concordo fortemente	Concordo	Não concordo
Psychological support (representa informações referente à assistência psicológica ao paciente presentes em um artigo científico)	○	○	○



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Classe: Pharmacologic Substance (Any substance, other than food, that is used to prevent, diagnose,

As propriedades a seguir obtiveram um percentual de 100% de aprovação:

- **Adverse effects**
- **Drug**
- **Biological effects**
- **Chemical name**
- **Dosage**
- **Mechanism of action**
- **Pharmacological action**
- **Reaction**

* 29. As propriedades a seguir obtiveram 88% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Brand name (representa o nome comercial da droga)	◯	◯	◯

Comente:

* 30. A propriedade a seguir obteve 63% de aprovação. Você concorda que elas devem sem mantidas?

	Concordo fortemente	Concordo	Não concordo
Radioiodine (refere-se ao uso do radioiodine)	◯	◯	◯

Deixe seu comentário:



MODELO DE DOMÍNIO PARA ARTIGOS CIENTÍFICOS EM ONCOLOGIA

Conclusão

Agradeço a sua participação neste estudo. Após o término desta investigação, gostaria de compartilhar os resultados, se for de seu interesse.

31. Informe o seu email, caso deseje receber a conclusão deste estudo.

APÊNDICE F – Esquema de metadados “ONCO” em RDF

DEFINIÇÃO DAS CLASSES

As classes foram definidas utilizando as propriedades do RDF e do RDF:

a) rdfs:Class

b) rdfs:subClassOf

c) rdfs:isDefinedBy

d) rdf:resource;

e) rdfs: comment; e

f) rdfs:label.

• FundingAgency

<rdfs:Class rdf:about=" http://opendata.dsi.uminho.pt/health/onco-schema#FundingAgency">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en"> FundingAgency </rdfs:label>
<rdfs:comment>Specific Class. A Funding Agency. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• Illustration

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#Illustration">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en"> Illustration </rdfs:label>
<rdfs:comment>Specific Class. A Illustration. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• ResearchDesign

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#ResearchDesign">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:subClassOf rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en"> ResearchDesign </rdfs:label>
<rdfs:comment>Specific Class. A Research Design. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• Neoplasm

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#Neoplasm">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=http://schema.org/ScholarlyArticle”/>
<rdfs:label xml:lang="en"> Neoplasm </rdfs:label>
<rdfs:comment>Specific Class. A Neoplasm. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• **Epidemiology**

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema/#Epidemiology">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=" http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">Epidemiology</rdfs:label>
<rdfs:comment>Specific Class. A Epidemiology. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• **Diagnosis**

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema/#Diagnosis">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=" http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">Diagnosis</rdfs:label>
<rdfs:comment>Specific Class. A Diagnosis. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• **Organism**

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema/#Organism">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=" http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">Organism</rdfs:label>
<rdfs:comment>Specific Class. A Organism. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• **Therapeutic**

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#Therapeutic">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=" http://schema.org/ScholarlyArticle "/>
<rdfs:label xml:lang="en">Therapeutic</rdfs:label>
<rdfs:comment>Specific Class. A Therapeutic. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• **PharmacologicSubstance**

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#PharmacologicSubstance">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=" http://schema.org/ScholarlyArticle "/>
<rdfs:label xml:lang="en">PharmacologicSubstance</rdfs:label>
<rdfs:comment>Specific Class. A Pharmacologic Substance. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

• **Anatomy**

<rdfs:Class rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#Anatomy">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:subClassOf rdf:resource=" http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">Anatomy</rdfs:label>
<rdfs:comment>Specific Class. A Anatomy. It is a subclass of ScholarlyArticle </rdfs:comment>
</rdfs:Class>

DEFINIÇÃO DAS PROPRIEDADES

As propriedades foram definidas utilizando as propriedades do RDF e do RDF:

a) rdf:Property

b) rdfs:isDefinedBy

c) rdfs:label;

d) rdfs: comment;

e) rdfs:domain; e

f) rdfs:range.

Foram criadas as seguintes propriedades:

- **score**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema#score">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://purl.org/ontology/bibo/Journal"-/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle”-/>
<rdfs:domain rdf:resource="http://purl.org/dc/terms/Agent”-/>
<rdfs:range rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">score</rdfs:label>
<rdfs:comment>Represents the metric for resource</rdfs:comment>
</rdf:RDF>

- **appendix**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# appendix">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">appendix</rdfs:label>
<rdfs:comment>Text or document prepared by the author</rdfs:comment>
</rdf:RDF>

- **illustration**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# illustration">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">illustration</rdfs:label>
<rdfs:comment>A image, graphic or table within a journal</rdfs:comment>
</rdf:RDF>

- **accessMode**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# accessMode ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">accessMode</rdfs:label>
<rdfs:comment>Declares the degree of openness of a resource</rdfs:comment>
</rdf:RDF>

- **fundingAgency**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# fundingAgency ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">fundingAgency</rdfs:label>
<rdfs:comment>Institutions that offer different forms of financing aimed at conducting research</rdfs:comment>
</rdf:RDF>

- **conflitofInterest**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# conflitOfInterest">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://schema.org/ScholarlyArticle"/>
<rdfs:label xml:lang="en">conflitOfInterest</rdfs:label>
<rdfs:comment>A situation in which an individual might benefit personally from oficial or professional actions.</rdfs:comment>
</rdf:RDF>

- **dataCollection**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# dataCollection">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# ResearchDesign"/>
<rdfs:label xml:lang="en">dataCollection</rdfs:label>
<rdfs:comment>Referes to process of gathering data for testing or research purposes.</rdfs:comment>
</rdf:RDF>

- **researchMethod**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# researchMethod ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# ResearchDesign"/>
<rdfs:label xml:lang="en">researchMethod</rdfs:label>
<rdfs:comment>A method or procedures used in a research</rdfs:comment>
</rdf:RDF>

- **statisticalTechnique**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# statisticalTechnique">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# ResearchDesign"/>
<rdfs:label xml:lang="en">statisticalTechnique</rdfs:label>
<rdfs:comment>A statistical technique used in a research</rdfs:comment>
</rdf:RDF>

- **sample**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# sample">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# ResearchDesign"/>
<rdfs:label xml:lang="en">sample</rdfs:label>
<rdfs:comment>The sample of study</rdfs:comment>
</rdf:RDF>

- **sampleSize**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# sampleSize">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# ResearchDesign"/>
<rdfs:label xml:lang="en">sampleSize</rdfs:label>
<rdfs:comment>The sample size</rdfs:comment>
</rdf:RDF>

- **classification**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# classification ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">classification</rdfs:label>
<rdfs:comment>The classification of cancers by histological type</rdfs:comment>
</rdf:RDF>

- **dateOfStage**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# dateOfStage">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">dateOfStage</rdfs:label>
<rdfs:comment>Date of he process of defining at what point in the natural history of a malignant disease a patient is when a diagnosis is made</rdfs:comment>
</rdf:RDF>

- **gradeTumor**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# gradeTumor ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">gradeTumor</rdfs:label>
<rdfs:comment>A position on a scale of intensity or amount or quality</rdfs:comment>
</rdf:RDF>

- **lesionCharacteristic**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# lesionCharacteristic">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">lesionCharacteristic</rdfs:label>
<rdfs:comment>Characteristic of localized pathological or traumatic structural change, damage, deformity, or discontinuity of tissue, organ, or body part</rdfs:comment>
</rdf:RDF>

- **metastasis**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# metastasis ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">metastasis</rdfs:label>
<rdfs:comment>The spread or migration of cancer cells from one part of the body to another</rdfs:comment>
</rdf:RDF>

- **recurrence**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# recurrence ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">recurrence</rdfs:label>
<rdfs:comment>represents the information If the cancer returned.</rdfs:comment>
</rdf:RDF>

- **site**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# site">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">site</rdfs:label>
<rdfs:comment>represents a position in which cancer develops in relation to morphomology</rdfs:comment>
</rdf:RDF>

- **size**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# size ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">size</rdfs:label>
<rdfs:comment> The physical magnitude of tumor</rdfs:comment>
</rdf:RDF>

- **statusAtStaging**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# statusAtStaging ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle"-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>
<rdfs:label xml:lang="en">statusAtStaging</rdfs:label>
<rdfs:comment> Level of extent of the cancer within the body</rdfs:comment>
</rdf:RDF>

- **causality**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# causality ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">causality</rdfs:label>
<rdfs:comment> Relationship exists between a cause or effect</rdfs:comment>
</rdf:RDF>

- **clinicalTrialInformation**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# clinicalTrialInformation ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">clinicalTrialInformation</rdfs:label>
<rdfs:comment> Report of a pre-planned medical or veterinary clinical case study involving a randomized, evidence-controlled clinical trial</rdfs:comment>
</rdf:RDF>

- **etiology**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# etiology ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">etiology</rdfs:label>
<rdfs:comment>The cause of a disease or abnormal condition</rdfs:comment>
</rdf:RDF>

- **geneticEpidemiology**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# geneticEpidemiology ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">geneticEpidemiology</rdfs:label>
<rdfs:comment>Interdisciplinary clinical, epidemiologic, genetic, and laboratory studies of persons, families, and populations at high risk of cancer.</rdfs:comment>
</rdf:RDF>

- **incidence**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# incidence ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">incidence</rdfs:label>
<rdfs:comment>The relative frequency of occurrence of cancer</rdfs:comment>
</rdf:RDF>

- **populationStatistics**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# populationStatistics ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">populationStatistics</rdfs:label>
<rdfs:comment>A value, which represents a particular population characteristic</rdfs:comment>
</rdf:RDF>

- **prevalence**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# prevalence ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">prevalence</rdfs:label>
<rdfs:comment>The total number of cases of a given disease in a specified population at a designated time</rdfs:comment>
</rdf:RDF>

- **primaryCauseofDeath**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# primaryCauseofDeath ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Epidemiology"/>
<rdfs:label xml:lang="en">primaryCauseofDeath</rdfs:label>
<rdfs:comment>The first significant event which ultimately led to death</rdfs:comment>
</rdf:RDF>

- **category**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# category ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis"/>
<rdfs:label xml:lang="en">category</rdfs:label>
<rdfs:comment>A category for the item</rdfs:comment>
</rdf:RDF>

- **healthCondition**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# healthCondition ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">healthCondition</rdfs:label>
<rdfs:comment>Specifying the health condition(s) of a patient, medical study, or other target audience</rdfs:comment>
</rdf:RDF>

- **interpretingClinicalFinding**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# interpretingClinicalFinding ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">interpretingClinicalFinding</rdfs:label>
<rdfs:comment>Interpretation of clinical finding</rdfs:comment>
</rdf:RDF>

- **criteriaManifestation**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# criteriaManifestation ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">criteriaManifestation</rdfs:label>
<rdfs:comment>Criteria based on patient symptoms and signs, test results</rdfs:comment>
</rdf:RDF>

- **pathologyFindings**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# pathologyFindings ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">pathologyFindings</rdfs:label>
<rdfs:comment>Pathology findings</rdfs:comment>
</rdf:RDF>

- **previousDiagnosis**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# previousDiagnosis ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">previous Diagnosis</rdfs:label>
<rdfs:comment>Previous diagnosis</rdfs:comment>
</rdf:RDF>

- **signOrSymptom**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# signOrSymptom ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">signOrSymptom</rdfs:label>
<rdfs:comment>A sign o r symptom of this condition</rdfs:comment>
</rdf:RDF>

- **vitalStatus**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# vitalStatus ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Diagnosis "/>
<rdfs:label xml:lang="en">vitalStatus</rdfs:label>
<rdfs:comment>The state or condition of being living or deceased; also includes the case where the vital status is unknown</rdfs:comment>
</rdf:RDF>

- **biomaker**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# biomaker ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema: DiagnosticProcedure"/>
<rdfs:label xml:lang="en">biomaker</rdfs:label>
<rdfs:comment>Biomarker used</rdfs:comment>
</rdf:RDF>

- **chemicalAgentAdministration**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# chemicalAgentAdministration ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://schema: DiagnosticProcedure"/>
<rdfs:label xml:lang="en">chemicalAgentAdministration</rdfs:label>
<rdfs:comment>Administration of chemical agent in exams</rdfs:comment>
</rdf:RDF>

- **diagnosticProcedure**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# diagnosticCriteria ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema: DiagnosticProcedure "/>
<rdfs:label xml:lang="en">diagnosticCriteria</rdfs:label>
<rdfs:comment>A standard from which a judgment concerning a medical diagnosis can be established</rdfs:comment>
</rdf:RDF>

- **stagingSystem**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# stagingSystem ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema: DiagnosticProcedure "/>
<rdfs:label xml:lang="en">stagingSystem</rdfs:label>
<rdfs:comment>A system adoped to describe the extent of cancer in a patient's body</rdfs:comment>
</rdf:RDF>

- **technique**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# technique ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# DiagnosticProcedure "/>
<rdfs:label xml:lang="en">technique</rdfs:label>
<rdfs:comment>Encompass all investigations and tests intended to identify the cause of disease</rdfs:comment>
</rdf:RDF>

- **test**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# test ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema: DiagnosticProcedure "/>
<rdfs:label xml:lang="en">test</rdfs:label>
<rdfs:comment>Process of determining the disease state of a person</rdfs:comment>
</rdf:RDF>

- **procedureType**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# procedureType ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema: DiagnosticProcedure"/>
<rdfs:label xml:lang="en">procedureType</rdfs:label>
<rdfs:comment>The type of procedure, for example surgical, noninvasive or percutaneous</rdfs:comment>
</rdf:RDF>

- **anatomicFrameOfReference**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# anatomicFrameOfReference ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">anatomicFrameOfReference</rdfs:label>
<rdfs:comment>Image acquisition module</rdfs:comment>
</rdf:RDF>

- **colorDepth**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# colorDepth ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">colorDepth</rdfs:label>
<rdfs:comment>Positive integer indicating the desired minimum total number of bits per color channel used to present a pixel</rdfs:comment>
</rdf:RDF>

- **magnification**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# magnification ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">magnification</rdfs:label>
<rdfs:comment>Magnification is related to scaling up visuals or images to be able to see more detail</rdfs:comment>
</rdf:RDF>

- **physicalDescription**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# physicalDescription ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">physicalDescription</rdfs:label>
<rdfs:comment>Physical description of the image</rdfs:comment>
</rdf:RDF>

- **primaryMagnification**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# primaryMagnification ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">primaryMagnification</rdfs:label>
<rdfs:comment>Initial magnification</rdfs:comment>
</rdf:RDF>

- **processingImage**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# processingImage ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://schema.org/ImagingTest"/>
<rdfs:label xml:lang="en">processingImage</rdfs:label>
<rdfs:comment>The use of data processing systems to analyze, enhance, interpret or display digital images</rdfs:comment>
</rdf:RDF>

- **sizeOfPixel**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# sizeOfPixel ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">sizeOfPixel</rdfs:label>
<rdfs:comment>Image size expressed in pixels</rdfs:comment>
</rdf:RDF>

- **software**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# software ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">software</rdfs:label>
<rdfs:comment>A computer program in source of compiled form</rdfs:comment>
</rdf:RDF>

- **visualizationMode**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# visualizationMode ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://schema.org/ImagingTest "/>
<rdfs:label xml:lang="en">visualizationMode</rdfs:label>
<rdfs:comment>Options for visualization</rdfs:comment>
</rdf:RDF>

- **proteinExpression**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# proteinExpression ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://dbpedia.org/ontology/Gene "/>
<rdfs:label xml:lang="en">proteinExpression</rdfs:label>
<rdfs:comment>The determination of the level of gene transcription in a biological specimen</rdfs:comment>
</rdf:RDF>

- **geneIdentifier**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# geneIdentifier ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Gene "/>
<rdfs:label xml:lang="en">geneIdentifier</rdfs:label>
<rdfs:comment>A unique gene name within a specific repository, database, or collection</rdfs:comment>
</rdf:RDF>

- **geneSymbol**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# geneSymbol ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://dbpedia.org/ontology/Gene "/>
<rdfs:label xml:lang="en">geneSymbol</rdfs:label>
<rdfs:comment>A unique gene name approved by an organism specific nomenclature committee</rdfs:comment>
</rdf:RDF>

- **cause**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# cause">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource=" http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">cause</rdfs:label>
<rdfs:comment>One of the causative agent that are most directly responsable for the pathophysiologic process that eventually results in the ocorrence</rdfs:comment>
</rdf:RDF>

- **characteristic**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# characteristic ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">characteristic</rdfs:label>
<rdfs:comment>The distinguishing of a disease aspect images</rdfs:comment>
</rdf:RDF>

- **classOfDisease**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# classOfDisease ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">classOfDisease</rdfs:label>
<rdfs:comment>Classification of diseases</rdfs:comment>
</rdf:RDF>

- **complication**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# complication ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">complication</rdfs:label>
<rdfs:comment>Any disease or disorder that occurs during the course of, or because of, another disease, treatment, or procedure</rdfs:comment>
</rdf:RDF>

- **diseaseAgent**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# diseaseAgent ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">visualizationMode</rdfs:label>
<rdfs:comment>Options for visualization</rdfs:comment>
</rdf:RDF>

- **visualizationMode**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# visualizationMode ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://schema.org/ImagingTest"/>
<rdfs:label xml:lang="en">diseaseAgent</rdfs:label>
<rdfs:comment>Any vírus, micro-organism, or other substance causing disease</rdfs:comment>
</rdf:RDF>

- **prognosis**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# prognosis ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">prognosis</rdfs:label>
<rdfs:comment>prognosis</rdfs:comment>
</rdf:RDF>

- **rareDisease**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# rareDisease ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">rareDisease</rdfs:label>
<rdfs:comment>Represents a rare disease</rdfs:comment>
</rdf:RDF>

- **adverseEffect**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# adverseEffect ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic"/>
<rdfs:label xml:lang="en">adverseEffect</rdfs:label>
<rdfs:comment>An unexpected medical problem that happens during treatment with a drug or other therapy</rdfs:comment>
</rdf:RDF>

- **contraindication**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# contraindication ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">contraindication</rdfs:label>
<rdfs:comment>A contraindication for this therapy</rdfs:comment>
</rdf:RDF>

- **currentTherapeutic**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# currentTherapeutic ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">currentTherapeutic</rdfs:label>
<rdfs:comment>The name of the therapy that is currently being administered</rdfs:comment>
</rdf:RDF>

- **drugClass**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# drugClass ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">drugClass</rdfs:label>
<rdfs:comment>The class of drug this belongs to</rdfs:comment>
</rdf:RDF>

- **efficacy**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# efficacy ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">efficacy</rdfs:label>
<rdfs:comment>The chemical agent or a device to treat a disease used successfuly</rdfs:comment>
</rdf:RDF>

- **indication**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# indication ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">indication</rdfs:label>
<rdfs:comment>A fator that indicates use of this therapy for treatment and/or prevention of a condition, symptom</rdfs:comment>
</rdf:RDF>

- **monitoring**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# monitoring ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">monitoring</rdfs:label>
<rdfs:comment>How a person or a condition is assited or verified.</rdfs:comment>
</rdf:RDF>

- **patientCompliance**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# patientCompliance ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">patientCompliance</rdfs:label>
<rdfs:comment>The act of following a medical regimen or schedule correctly and consistently, including taking medicines or following a diet rdfs:comment>
</rdf:RDF>

- **prevention**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# prevention ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">prevention</rdfs:label>
<rdfs:comment>Describes methods of preventing cancer including limit of exposure to carcinogens and infectious agents, diet, education rdfs:comment>
</rdf:RDF>

- **medicalProcedure**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# medicalProcedure ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">medicalProcedure</rdfs:label>
<rdfs:comment>Describes a treatment or action taken to prevent or treat disease
<rdfs:comment>
</rdf:RDF>

- **protocol**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# protocol ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">protocol</rdfs:label>
<rdfs:comment>Describes a rule which guides how an activity should be performed<rdfs:comment>
</rdf:RDF>

- **typeOfTreatment**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# typeOfTreatment ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">typeOfTreatment</rdfs:label>
<rdfs:comment>Cancer treatment adopted to non directed or directed cancer <rdfs:comment>
</rdf:RDF>

- **psychologicalAssessment**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# psychologicalAssessment">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Therapeutic "/>
<rdfs:label xml:lang="en">psychologicalAssessment</rdfs:label>
<rdfs:comment> Represents information regarding the psychological assistance to the patient <rdfs:comment>
</rdf:RDF>

- **adverseEffect**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# adverseEffect ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">adverseEffect</rdfs:label>
<rdfs:comment>An unexpected medical problem that happens during treatment with a drug or other
therapy<rdfs:comment>
</rdf:RDF>

- **drug**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# drug ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">drug</rdfs:label>
<rdfs:comment>Specifying a drug or medicine used in a medication procedure<rdfs:comment>
</rdf:RDF>

- **biologicalEffect**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# biologicalEffect ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">biologicalEffect</rdfs:label>
<rdfs:comment>Biological effect that the any product provide<rdfs:comment>
</rdf:RDF>

- **chemicalName**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# chemicalName ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">chemicalName</rdfs:label>
<rdfs:comment>Chemical name. <rdfs:comment>
</rdf:RDF>

- **pharmacologicAction**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicAction ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">pharmacologicAction</rdfs:label>
<rdfs:comment>Actions of drugs and their effects on living organisms<rdfs:comment>
</rdf:RDF>

- **radioiodine**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# radioiodine ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">radioiodine</rdfs:label>
<rdfs:comment>Use of radioiodine<rdfs:comment>
</rdf:RDF>

- **reaction**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# reaction ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# pharmacologicSubstance "/>
<rdfs:label xml:lang="en">reaction</rdfs:label>
<rdfs:comment>A description of the action in response to some kind of stimulus<rdfs:comment>
</rdf:RDF>

- **anatomicRegion**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# anatomicRegion ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#"/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Anatomy"/>

<rdfs:label xml:lang="en">anatomicRegion</rdfs:label>
<rdfs:comment>Identifies areas of the body<rdfs:comment>
</rdf:RDF>

- **biologicalProcess**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# biologicalProcess ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Anatomy"/>
<rdfs:label xml:lang="en">biologicalProcess</rdfs:label>
<rdfs:comment>An activity occurring within an organism, between organisms or among organisms and the mechanisms underlying such events<rdfs:comment>
</rdf:RDF>

- **cellComponen**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# cellComponen">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Anatomy"/>
<rdfs:label xml:lang="en">cellComponen</rdfs:label>
<rdfs:comment>The components of the cell<rdfs:comment>
</rdf:RDF>

- **proteinStructure**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# proteinStructure ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Anatomy"/>
<rdfs:label xml:lang="en">proteinStructure</rdfs:label>
<rdfs:comment>The amino acids and their 3-D arrangement in constituting a protein<rdfs:comment>
</rdf:RDF>

- **tissueType**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# tissueType ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Anatomy"/>
<rdfs:label xml:lang="en">tissueType</rdfs:label>
<rdfs:comment>Describes collections of differentiated cells, such as epithelium; connective tissue; muscles; nerve tissue<rdfs:comment>
</rdf:RDF>

- **Mechanism biologic**

<rdf:Property rdf:about="http://opendata.dsi.uminho.pt/health/onco-schema# tissueType ">
<rdfs:isDefinedBy rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema#”/>
<rdfs:domain rdf:resource="http://schema.org/ScholarlyArticle "-/>
<rdfs:range rdf:resource="http://opendata.dsi.uminho.pt/health/onco-schema# Neoplasm"/>

<rdfs:range rdf:resource="http://dbpedia.org/ontology/Disease "/>
<rdfs:label xml:lang="en">biologicMechanism</rdfs:label>
<rdfs:comment> Biological mechanism associated with disease<rdfs:comment>
</rdf:RDF>