

Uma Ontologia para Transparência no Contexto do Governo Aberto

A Transparency Ontology in the Context of Open Government

Rui Pedro Lourenço, INESC Coimbra e Faculdade de Economia da Universidade de Coimbra, Portugal, ruiloure@fe.uc.pt

Helena Rolas, INESC Coimbra e Faculdade de Economia da Universidade de Coimbra, Portugal, lenarolas@hotmail.com

Susana Jorge, NEAPP (Universidade do Minho) e Faculdade de Economia da Universidade de Coimbra, Portugal, susjor@fe.uc.pt

Resumo

No contexto do governo aberto surgiram diversas iniciativas visando a disponibilização online de dados sobre o funcionamento das entidades públicas, de forma a facilitar prestação de contas (transparência) e a responsabilização dos agentes públicos por parte dos cidadãos (accountability). Estas iniciativas levantam, no entanto, diversas questões, entre as quais: que dados disponibilizar? Como tornar coerentes os dados dispersos (fragmentados) por diversos portais? Como tornar os dados disponibilizados mais compreensíveis para o cidadão comum? Este trabalho tem como objetivo definir uma ontologia para dar resposta a estas questões, colmatando assim uma lacuna identificada nesta área. A ontologia proposta procura adotar a perspetiva do cidadão comum, partindo da identificação de temas/termos relevantes debatidos na comunicação social (escrita), considerando o papel dos jornalistas enquanto information brokers. Espera-se que através da ontologia aqui apresentada se possa melhorar a forma como as estratégias de governo aberto são conduzidas e os portais que dela resultam.

Palavras chave: Governo aberto, transparência, accountability, ontologia

Abstract

Several initiatives have been launched, within the context of open government, to make available online data concerning the functioning of public entities. Such initiatives aim to improve transparency of public entities and to enhance accountability of public officials towards ordinary citizens. However, these initiatives raise several questions, namely: Which data should be disclosed? How to bring together dispersed (fragmented) data disclosed through several different portals? How to improve the understandability by ordinary citizens of the data disclosed? The goal of this paper is to define and propose a transparency ontology that might help to provide an answer to these questions, thus filling the gaps in this research area. The proposed ontology seeks to adopt the perspective of the ordinary citizen in what concerns public entities' transparency and accountability. Therefore, the process of ontology definition started by identifying relevant themes/terms discussed in the media (newspapers) considering the role of journalists as 'information brokers'. The proposed ontology might be used by public officials to improve the formulation of open government strategies and the way they are implemented through open government data portals.

Keywords: Open Government, transparency, accountability, ontology

1. INTRODUÇÃO

Nos últimos anos o conceito de *governo aberto* (*open government*), nas suas múltiplas vertentes, tem aumentado a sua visibilidade e importância na esfera pública e na agenda política. Um impulso importante para a afirmação da sua relevância veio sem dúvida do *Memorandum on Transparency and Open Government*, emitido pelo Presidente Barack Obama em 2009 [Obama, 2009], consubstanciado em seguida pela *Open Government Directive* [Orszag, 2009], que em traços gerais, procurava promover um governo mais transparente, participativo e colaborativo. Este primeiro ímpeto deu início a diversas outras iniciativas, tais como a *Open Government Partnership*¹ que envolve neste momento 64 países comprometidos em “promover a transparência, dar poder aos cidadãos, combater a corrupção e aproveitar o potencial das novas tecnologias para fortalecer a governança”².

De entre os diversos objetivos associados ao *governo aberto* [Linders e Wilson, 2011], é o objetivo da *transparência* (disponibilização de dados das Administrações Públicas, *Public Sector Information* – PSI) aquele que parece neste momento ter uma maior tradução prática, nomeadamente com o surgimento de diversos portais. Surge assim o Data.Gov, porventura o mais prominente desses portais, que é apenas um entre muitos outros (temáticos ou genéricos) promovidos pela Administração Norte-Americana³, e onde é possível encontrar uma lista de cerca de 300 outros portais deste tipo (internacionais, estaduais e de cidades americanas⁴).

Para além do valor que os dados disponibilizados possam ter ao ser re-utilizados para a criação de novos produtos e serviços, a sua disponibilização através de portais pode desempenhar um papel crucial no sistema democrático, ao permitir que os cidadãos saibam como funcionam as entidades públicas e como são geridos os recursos colocados à sua disposição. Neste contexto, os portais governamentais de dados abertos constituem ferramentas importantes de promoção da transparência (prestação de contas) visando a responsabilização dos agentes públicos (*accountability* [Bovens, 2005; Bovens, 2007]).

O desenvolvimento destes portais tem sido, mais recentemente, associado a alguns conceitos fundamentais que o portal Data.gov.uk descreve no seu glossário: *Linked Data*⁵, *Open Data*⁶,

¹ <http://www.opengovpartnership.org/>

² <http://www.opengovpartnership.org/about>

³ <http://www.whitehouse.gov/open/about/initiatives>

⁴ <http://www.data.gov/open-gov/>

⁵ “The term used to describe the recommended best practice for exposing, sharing and connecting items of data on the semantic web using unique resource identifiers (URIs) and resource description framework (RDF).” <http://data.gov.uk/glossary/23179>

⁶ “Data which can be used, re-used and re-distributed freely by anyone - subject only at most to the requirement to attribute and share-alike. There may be some charge, usually no more than the cost of reproduction.” <http://data.gov.uk/glossary/23187>

*Semantic Web*⁷, ou Ontologia⁸. É precisamente sobre este último que o presente trabalho incide, procurando contribuir para colmatar duas lacunas identificadas no âmbito das iniciativas de *governo aberto*, nomeadamente nas que têm por objetivo fomentar a transparência que visa a *accountability* dos agentes públicos através da disponibilização *online* de dados relevantes:

- Que tipo de dados (itens de informação) devem as entidades públicas disponibilizar (*online*) de forma a melhor corresponder aos tópicos que os cidadãos, últimos destinatários da transparência e principais partes interessadas na responsabilização dos agentes públicos, consideram relevantes?
- Como estruturar a forma fragmentada como são disponibilizados dados, quer através de inúmeros portais temáticos (como, p. ex., o Portal Base⁹), quer através de diferentes *datasets* em portais globais (como, p. ex., Dados.gov¹⁰)?

No que diz respeito ao primeiro ponto, existem já diversos modelos de avaliação da informação disponibilizada na internet pelas entidades públicas (veja-se, por exemplo [Lourenço, 2013a]) que definem geralmente um conjunto de itens de informação que devem ser disponibilizados *online* pelas entidades avaliadas. No entanto, não foi encontrado nenhum modelo em que a escolha dos itens resultasse da consideração do ponto de vista dos cidadãos (destinatários da informação). Em geral, os itens são selecionados a partir de requisitos legais ou normativos respeitantes a informação a divulgar pelas entidades públicas, ou simplesmente incluindo documentos tradicionais de prestação de contas (orçamentos, planos de atividades, ...). Por sua vez, é invocada, geralmente de forma implícita, a qualidade de perito dos autores do estudo para justificar a escolha dos itens e a sua relevância.

Esta é portanto uma das lacunas a que este trabalho pretende dar resposta, propondo um conjunto de itens de informação (definidos através da ontologia proposta) cuja escolha teve como ponto de partida o ponto de vista dos cidadãos. No entanto, na abordagem seguida optou-se por não questionar diretamente cidadãos comuns (através de questionários, por exemplo) por se considerar que nem sempre estes possuem o conhecimento necessário ou o interesse em analisar diretamente as fontes primárias de informação colocadas à sua disposição, criando assim o que Heald designa por “problema da ausência de utilizadores” [Heald, 2003; Heald, 2012]. Considerando então que os

⁷ “A web of data that can be processed directly and indirectly by machines, providing a common framework that allows data to be shared and reused across application, enterprise, and community boundaries. It is based on the Resource Description Framework (RDF).” <http://data.gov.uk/glossary/23200>

⁸ “Formal representation of knowledge as a set of concepts within a domain, and the relationships among those concepts.” <http://data.gov.uk/glossary/23185>

⁹ Portal dos Contratos Públicos (www.base.gov.pt): “Aqui, é comunicada a informação sobre todos os contratos celebrados ao abrigo do Código dos Contratos Públicos.”

¹⁰ No caso português o portal Dados.gov tem o endereço <http://dados.gov.pt>

cidadãos dependem de *intermediários informacionais* [Fung, 2013] ou *information brokers*¹¹, como é o caso dos jornalistas, optou-se pela análise de um mês de notícias publicadas em dois jornais de circulação nacional e um local. As notícias selecionadas para análise visavam aspetos em geral da gestão de entidades públicas, sendo consideradas de interesse para os cidadãos pelos jornalistas.

No que diz respeito ao segundo ponto, a definição de uma ontologia neste domínio (transparência para *accountability*) pode dar resposta ao problema da fragmentação e falta de ligação dos dados disponibilizados através de diferentes portais, bem como facilitar a pesquisa e a compreensão do significado de alguns termos empregues neste contexto. Embora existam já diversas ontologias relacionadas com governo eletrónico, estas abrangem tópicos específicos tais como a integração de serviços públicos [Vitvar e Nazir, 2007], aspetos particulares da prestação de contas, tais como *contratos públicos* [Klímek *et al.*, 2012] ou *pagamentos* (*Payments Ontology*¹²), ou são dedicadas a uma administração pública em particular (USA – oeGOV¹³). Neste contexto não foi encontrada nenhuma que vise, de uma forma genérica, a transparência para *accountability* nas administrações públicas.

No caso concreto da realidade portuguesa, foi já identificado um conjunto de portais de gestão/disponibilização de dados relevantes no contexto da responsabilização dos agentes públicos pela gestão dos recursos de que dispõem [Lourenço, 2013b]. No entanto, existe a perceção de que a criação dos mesmos não obedece a uma lógica de coerência global, resultando numa fragmentação tal que dificulta a integração fácil dos dados dispersos para a realização de análises que exijam a consulta de diferentes fontes.

A título de exemplo, dois dos portais mencionados no estudo sobre a realidade portuguesa (SIOE e Portal Base) [Lourenço, 2013b] definem de forma diferente o universo de ‘entidades públicas’ abrangidas¹⁴, recorrendo a uma linguagem (em certa medida) técnica. Considerando que cabe ao SIOE o papel definidor e caracterizador do universo de ‘entidades públicas’, a diferença nas definições apresentadas torna impossível verificar se a não existência de dados no Portal BASE acerca de uma entidade que consta do SIOE se deve à sua falta de cumprimento dos requisitos legais, ou simplesmente ao facto de os universos abrangidos serem diferentes. O cidadão comum, isto é, não especialista em matérias técnicas relacionadas com a organização do Estado, vê-se assim

¹¹ “... que Rutherford [1992] descreve como ‘utilizadores intermédios’ de relatórios financeiros governamentais e documentos similares.” [Heald, 2003]

¹² Definida pelo Governo do Reino Unido no âmbito do seu programa de Governo Aberto (<http://data.gov.uk/resources/payments>)

¹³ <http://oegov.org/>

¹⁴ SIOE - <http://www.sioe.dgaep.gov.pt/Default.aspx>;

Portal BASE - <http://www.base.gov.pt/base2/html/utilidades/faq.shtml>

impossibilitado de realizar um simples juízo acerca da atuação dos responsáveis pelas entidades públicas em causa.

Este pequeno exemplo ilustra, de forma cabal, o facto de que não é suficiente para uma estratégia de governo aberto a simples disponibilização de dados em portais dedicados às diferentes temáticas. É importante que a escolha dos dados a disponibilizar vá de encontro, tanto quanto possível, às necessidades dos cidadãos no que diz respeito à responsabilização dos agentes públicos, facilitando simultaneamente a compreensão do que é disponibilizado, e possibilitando análises que envolvam a conjugação de dados disponibilizados através de diferentes portais. É esse o papel crucial que uma ontologia dedicada à transparência para *accountability* pode desempenhar, e é essa a contribuição que este trabalho procura dar.

A fim de concretizar estes objetivos, o trabalho encontra-se estruturado da seguinte forma. Na secção 2 serão apresentados alguns conceitos que facilitam a compreensão do âmbito da ontologia desenvolvida seguindo-se, na secção 3, uma breve explicação dos principais aspetos metodológicos que enquadraram o trabalho. Na secção 4 será depois apresentada a ontologia desenvolvida, nas suas diferentes etapas, terminando o artigo com algumas considerações finais.

2. ONTOLOGIAS

Segundo Gruber [1993], uma ontologia pode ser definida como “*a specification of a representational vocabulary for a shared domain of discourse*”. Sobre o que constitui uma ontologia, Uschold e Gruninger [2004] sugerem que o conceito pode abranger um espectro contínuo que vai desde um (simples) glossário ou lista de termos, até à utilização de lógica formal para a definição de regras semânticas do vocabulário. Numa posição intermédia, apresentam como exemplo os modelos conceptuais de entidades e relacionamentos (modelos E/R) ou esquemas de implementação de bases de dados. Assim, no que diz respeito à formalidade de uma ontologia, é possível distinguir quatro graus (pontos arbitrários ao longo do espectro contínuo) [Uschold e Gruninger, 1996]: ontologias de elevada informalidade (*highly informal* - expressas de forma livre em linguagem natural); ontologias semi-informais (expressas em linguagem natural mas de forma estruturada para assim reduzir a ambiguidade e aumentar a clareza); ontologias semi-formais (expressas numa linguagem artificial); ontologias rigorosamente formais (expressas numa terminologia definida através de regras semânticas, e incluindo teoremas e demonstrações formais).

Para além do grau de formalidade, as ontologias podem também ser caracterizadas pela sua generalidade, incluindo [Studer *et al.*, 1998]: ontologias de domínio (cujo objetivo é capturar o conhecimento num determinado domínio do saber); ontologias genéricas (aplicáveis a diferentes domínios e também conhecidas como *core ontologies*); ontologias de aplicação (abrangem todo o

conhecimento necessário para modelar um determinado domínio, incluindo os métodos nele utilizados); e ontologias representativas (definem entidades representativas abstratas, sem definir o que deve ser representado em concreto).

Independentemente destas possíveis variações, Uschold e Gruninger [2004] identificam duas características essenciais comuns a qualquer ontologia: um vocabulário (lista) de termos que se referem a *coisas (things)* relevantes num determinado domínio; e uma especificação sobre o significado desses termos. Relativamente a esta última característica referem que, idealmente, essa especificação deve obedecer a uma qualquer lógica.

A partir destas duas dimensões de caracterização podemos afirmar que a ontologia proposta neste trabalho é uma *ontologia de domínio*, com uma definição *semi-formal*. Com efeito, a ontologia proposta visa precisamente a definição de um vocabulário comum que represente o domínio discursivo (na esfera pública) relacionado com a transparência pública e a responsabilização dos agentes das entidades das Administrações Públicas (*accountability*).

Já o grau de formalidade adotado neste trabalho para especificar a ontologia proposta resulta, em parte, do facto de o ponto de partida para o seu desenvolvimento ser constituído por notícias publicadas (em linguagem natural, portanto) em jornais. Procurou-se assim fundamentar a escolha da informação relevante na perspetiva dos cidadãos.

3. METODOLOGIA

Em termos genéricos a investigação realizada pode ser enquadrada num paradigma de *Design Science Research (DSR)* [Hevner *et al.*, 2004; Peffers *et al.*, 2007], na medida em que se pretendeu desenvolver um artefacto (uma ontologia) de forma a responder a um problema relevante: a falta de uma macro-estrutura que possibilite a integração, pesquisa e navegação de dados de *accountability* disponibilizados de forma fragmentada através de diferentes portais, servindo ainda de referencial conceptual comum para definir os itens de informação a disponibilizar e facilitar a sua compreensão e análise.

O processo de desenvolvimento da ontologia seguiu, em traços gerais, a proposta de Noy e McGuinness [2001], segundo duas grandes fases: *especificação* e *conceptualização* [Brusa *et al.*, 2006].

No que diz respeito à fase de *especificação*, começou-se por determinar o objetivo e âmbito da ontologia (já mencionados), a partir de uma descrição do domínio (transparência que visa a responsabilização – ver secção 1). Foi ainda definida a granularidade e o tipo de ontologia pretendido, bem como o seu grau de formalismo. A fase de *conceptualização* englobou a enumeração dos termos mais importantes do domínio (passo 3 de Noy e McGuinness [2001]), bem

como a sua definição, hierarquização e relacionamento (passos 4 e parte do passo 5 de Noy e McGuinness [2001]). Dado que não era objetivo deste trabalho, pelo menos nesta primeira abordagem, o detalhe e formalização completa da ontologia com vista à sua implementação imediata, não foram realizados os passos seguintes na fase de especificação, parando-se então no semi-formalismo da sua especificação.

Apesar desta ausência de avaliação externa formal, o facto de terem sido seguidos os passos propostos na literatura para o desenvolvimento de ontologias [Noy e McGuinness, 2001; Brusa *et al.*, 2006] proporcionou o rigor na investigação segundo as linhas de orientação de DSR [Hevner *et al.*, 2004].

O processo de desenvolvimento da ontologia passou por quatro etapas. A primeira etapa consistiu na recolha de notícias publicadas em dois jornais de circulação nacional e um de circulação local, considerando o papel de *information brokers* desempenhado pelos jornalistas. Foram seleccionadas separadamente (pelos investigadores da equipa) notícias publicadas ao longo de um mês que, de alguma forma, refletissem tópicos suscetíveis de ser considerados no âmbito da prestação de contas e responsabilização dos agentes públicos por parte dos cidadãos.

A segunda etapa consistiu na análise dos extratos de notícias recolhidos, visando a identificação e definição dos termos mais importantes do domínio da transparência para *accountability* (passo 3 de Noy e McGuinness [2001]). Para isso recorreu-se à técnica de análise qualitativa de conteúdo [Hsieh e Shannon, 2005].

Numa terceira etapa procedeu-se à elaboração de frases que refletissem, de forma abstrata, os excertos e títulos das notícias seleccionadas recorrendo aos termos identificados. Este exercício permitiu identificar, de forma clara, as relações possíveis entre os termos. Pode dizer-se que o resultado desta fase foi uma primeira ontologia semi-informal, na medida em que as frases abstratas foram expressas em linguagem natural (utilizando os termos listados na primeira etapa), mas de uma forma estruturada para aumentar a sua clareza.

Finalmente, numa quarta etapa, as frases abstratas foram representadas graficamente (aumentando assim o grau de formalismo e a clareza da ontologia), identificando claramente os conceitos principais e a terminologia do domínio, a forma como estes conceitos se relacionam, e criando um ponto de partida para acesso às definições entretanto produzidas.

A secção seguinte apresenta o processo de desenvolvimento da ontologia final, bem como exemplos das versões intermédias que foram sendo produzidas (com graus de formalismo crescente).

4. A ONTOLOGIA PROPOSTA: PROCESSO DE DESENVOLVIMENTO E EXEMPLOS

4.1. *Análise de notícias e recolha de extratos*

O ponto de partida para o desenvolvimento da ontologia foi, tal como mencionado anteriormente, a análise de notícias publicadas em três jornais durante o mês de janeiro de 2014: dois jornais nacionais (*Público* e *Jornal I*) e um jornal de circulação local (*Diário de Coimbra*). Relativamente aos jornais nacionais, de entre os que se podem considerar como sendo generalistas, estes foram considerados como tendo grande visibilidade e apresentarem um maior foco na divulgação de notícias relacionadas com as áreas que se pretendiam investigar. Para além disso, na escolha pesou também a acessibilidade da equipa de investigadores a estas duas publicações, quer em suporte físico, quer *online*. No que diz respeito ao *Diário de Coimbra*, a escolha por este jornal prendeu-se, sobretudo, com fatores de conveniência, dado que é o jornal da área geográfica de trabalho dos investigadores, o que, mais uma vez, garantiu um fácil acesso e consulta.

Nesta etapa, os investigadores consultaram os jornais de forma independente, identificando notícias relevantes e extraíndo de cada notícia o seu título e frases significativas (extratos). Foram definidas algumas diretrizes para orientação de todos no processo de recolha. Após este trabalho individual dos investigadores procedeu-se à consolidação dos extratos recolhidos numa lista única. Dadas as orientações que cada um seguiu para a recolha, de forma geral as notícias selecionadas por cada investigador foram sensivelmente as mesmas. Para os casos em que tal não aconteceu tentou-se refletir em conjunto sobre a relevância ou não da inclusão de alguns dos extratos.

No total, foram recolhidas 267 notícias (71 do *Público*, 64 do *Jornal I* e 132 do *Diário de Coimbra*).

O extrato apresentado de seguida ilustra uma notícia recolhida associada ao tema do endividamento, designadamente referindo dívidas a fornecedores e empréstimos obtidos.

Título: “CDU pede à Inspeção-Geral de Finanças auditoria às contas da Câmara de Gaia”

Extrato: “Dívidas a fornecedores, ..., um empréstimo no valor de 13 milhões de euros, 4,7 milhões de euros de créditos do município.”

Jornal: Público, edição de 24/01/2014

Este extrato exemplo será utilizado ao longo desta secção para ilustrar todo o processo de definição da ontologia proposta.

4.2. *Lista de expressões/termos relevantes*

Numa segunda fase procedeu-se à análise qualitativa de conteúdo [Hsieh e Shannon, 2005] da lista conjunta (unificada) de extratos. Procurou-se desta forma associar a cada extrato as

expressões/termos mais relevantes representativos do seu conteúdo, numa perspetiva de divulgação de aspetos concretos da gestão dos recursos (materiais, financeiros e humanos). As expressões/termos dos diversos extratos foram depois, por sua vez, também consolidados numa lista final organizada segundo temas comuns (Tabela 1).

Temas	Expressões/termos
Documento (tipos)	De suporte a ato de gestão; Plano ou relatório
Agente (tipos)	Entidade pública; Entidade privada; Indivíduo
Entidades públicas (eventos)	Criação; Extinção; Realização de ação de formação
Receita (eventos)	Transferência (recebida); Subsídio; Taxa; Tarifa; Imposto
Despesa (eventos)	Transferência (efetuada); Subsídio, apoio ou subvenção; Indemnização; Extraordinária
Endividamento (eventos)	Empréstimo (obtido); Empréstimo (concedido); Dívida a fornecedor
Contratação (eventos)	Avença
	Aquisição de bem móvel ¹⁵
	Contratação de serviço
	Contrato de empreitada
Património (tipos)	Imóvel; Terreno
Património (eventos)	Concessão; Cedência; Arrendamento (a); Arrendamento (de); Expropriação; Permuta; Aquisição; Venda; Empreitada/obra
Investimento (eventos)	Bem móvel (tipo) ¹⁶ ; Participação social; Parceria Público-Privada (PPP)
Comissões/Grupos de Trabalho (eventos)	Criação; Extinção
Litígios judiciais (eventos)	Processo instaurado; Processo de que foi alvo

¹⁵ Respeitante ao processo de aquisição de um determinado bem.

¹⁶ Respeitante ao valor investido (total) num determinado tipo de bens móveis.

Temas	Expressões/termos
Pessoal (eventos)	Contratação; Rescisão; Processo disciplinar; Transferência (de); Transferência (para)
Cargos/Órgãos (eventos)	Nomeação; Demissão; Cessação de funções

Tabela 1 – Lista de temas e respetivos termos

No processo de elaboração da lista, para além de se procurar uniformizar as expressões/termos (p. ex., imóvel = edifício), houve a preocupação de encontrar um nível de detalhe consistente e adequado ao objetivo inicial através da identificação de relações de generalização/especialização entre as expressões/termos identificados, bem como de os enquadrar em temas comuns. Estes são meramente indicativos, não tendo a agregação sugerida a pretensão de ser a única forma de os agrupar.

É importante referir também que nem todas as expressões/termos da Tabela 1 resultaram diretamente da análise dos extratos das notícias. Por exemplo, embora nos extratos não fosse encontrado o termo *indivíduo*, ele surge aqui listado porque se constatou que em algumas notícias eram referidos indivíduos específicos como agentes ou alvos de alguma atividade reportada.

É ainda de realçar que a lista de termos apresentada na Tabela 1 não resultou imediata e unicamente da etapa anterior, tendo sido refinada, complementada e corrigida à medida que decorreu o trabalho nas etapas seguintes.

Para cada expressão/termo listado iniciou-se depois um processo de definição do seu significado (conceito) partindo, sempre que possível de fontes oficiais, incluindo legislação. Quando não foram encontradas definições que se enquadrassem nos objetivos e âmbito da ontologia proposta (utilizando uma linguagem demasiado técnica ou orientada para outras áreas), foram propostas definições pelos próprios investigadores.

4.3. Abstração e ontologia semi-informal

Na fase seguinte procurou-se analisar as relações possíveis entre as diversas expressões/termos identificados, definindo para isso um conjunto de frases abstratas, utilizando linguagem natural de forma estruturada. Pretendia-se que estas frases refletissem as ações sujeitas ao escrutínio público relatadas no conjunto de extratos identificados. As frases recorrem à terminologia identificada na Tabela 1, podendo a mesma expressão/termo assumir diferentes papéis gramaticais. Estas frases, ao relacionarem as diferentes expressões/conceitos, serviram de guia para a definição gráfica da ontologia apresentada na secção seguinte.

De uma forma geral foram criadas frases correspondentes às seguintes tipologias:

- Frases representando ações sem sujeito definido
Exemplo: “*Foi criada a entidade pública A*” ou “*Foi extinta a entidade pública B*”

Em todas as restantes tipologias presume-se que o sujeito é uma entidade pública específica (que é o “objeto da transparência”).

- Frases representando ações sem agente destinatário¹⁷ (entidade pública, entidade privada ou indivíduo) definido, nem envolvendo outros objetos¹⁸ (imóveis, terrenos, ...)
Exemplo: “*A entidade pública A realizou uma ação de formação*”

Neste exemplo não se pretende especificar quem participou na formação, apenas que ela se realizou e teve um dado custo, um objetivo, abrangeu uma dada temática, ... (entre outros atributos possíveis).

- Frases representando ações sem agente destinatário (entidade pública, entidade privada ou indivíduo) definido, mas envolvendo outros objetos (imóveis, terrenos, ...)
Exemplo: “*A entidade pública A gastou x euros em obras no terreno/imóvel Z*”
- Frases representando o efeito agregado de várias ações, sem discriminar as ações isoladas abrangidas
Exemplo: “*A entidade pública A recebeu x euros, no total, em taxas do tipo T*”

Nesta tipologia considerou-se que a disponibilização de dados sobre todas as ações individuais (associadas a cada pagamento de cada tipo de taxa, por exemplo) representaria um nível de detalhe que, para além de não se justificar, geraria também um volume de dados considerável. Assim, opta-se por identificar/enumerar *tipos* a partir dos quais os dados são agregados. Outros exemplos desta tipologia dizem respeito a *Tarifas e Impostos* (do tema *Receitas*), a *Despesas Extraordinárias* ou a investimento total em diferentes tipos de *Bens Móveis*.

- A tipologia mais comum diz respeito a frases que refletem o reportar de uma ação isolada, realizada por uma entidade pública alvo de escrutínio, que incide sobre um dos três tipos de agentes destinatários reconhecidos (entidades públicas, privadas ou indivíduos).
Exemplo: “*A Entidade Pública A estabeleceu um contrato de empreitada com a Entidade Privada B, no valor de x €*”
- Finalmente, uma outra tipologia estende a tipologia anterior considerando que algumas ações podem ainda envolver outros objetos.
Exemplo: “*A entidade pública A concessionou o terreno/imóvel Z à entidade privada/indivíduo B*”

¹⁷ A expressão *agente destinatário*, aqui adotada, pretende designar um agente (entidade pública, entidade privada ou indivíduo) que surja como complemento direto ou indireto no predicado da frase

¹⁸ O termo *objeto*, aqui adotado, pretende designar um tipo de documento ou património que surja como complemento direto ou indireto no predicado da frase

Nestas duas últimas tipologias é necessário analisar, para cada ação, que tipo(s) de agente(s) podem estar envolvidos, isto é, por exemplo, se uma entidade pública pode ou não estabelecer um contrato de empreitada com um indivíduo, ou se pode concessionar um terreno a outra entidade pública.

Relativamente ao exemplo apresentado no início da secção, relacionado com o endividamento, a frase abstrata correspondente ao extrato de jornal faz parte da última tipologia identificada:

“A entidade pública A contraiu um empréstimo no valor de x euros junto da entidade pública / entidade privada B”

Note-se que para além do valor do empréstimo, muitos outros atributos poderão ser considerados (p. ex., data de vencimento, periodicidade da prestação, montante já reembolsado, ...). A definição do detalhe a divulgar sobre cada item será efetuada numa fase posterior do desenvolvimento da ontologia, não apresentada neste artigo.

A tipificação das frases abstratas, no sentido de identificar as relações entre as diversas expressões/termos e temas, permitiu ainda identificar um caso particular: o dos *Documentos*. Em muitos casos os extratos mencionavam a existência de documentos associados aos mais diversos atos de gestão (contratos, pareceres, relatórios produzidos pelas comissões ou grupos de trabalho criados, ...) ou decorrentes da atividade regular da entidade pública (no caso das Câmaras Municipais, por exemplo, podem citar-se os Planos Diretores Municipais, Planos de Ordenamento, Orçamentos, Relatórios de Gestão e Contas, ...). Assim, dada a enorme variedade de situações em que é possível associar um documento a uma ação, estes não foram considerados na formulação das frases abstratas deixando em aberto a possibilidade de se relacionarem com qualquer tipo de agente ou ação por ele exercida.

4.4. Representação gráfica e ontologia semi-formal

O passo seguinte na formalização da ontologia passou pela sua representação gráfica. Também aqui se optou por, nesta fase, não adotar uma simbologia excessivamente formal (como o UML – *Unified Modeling Language*). Em vez disso adotou-se uma simbologia simples tal como a que é utilizada na ontologia *Payments* (e outras) já referida, tendo apenas por objetivo (nesta fase) tornar claro que tipos de dados deverão ser alvo do processo de transparência, bem como algumas das possíveis relações entre eles.

A Figura 1 apresenta uma parte do diagrama elaborado, nomeadamente focando o exemplo que tem vindo a ser acompanhado ao longo do presente artigo (*endividamento*).

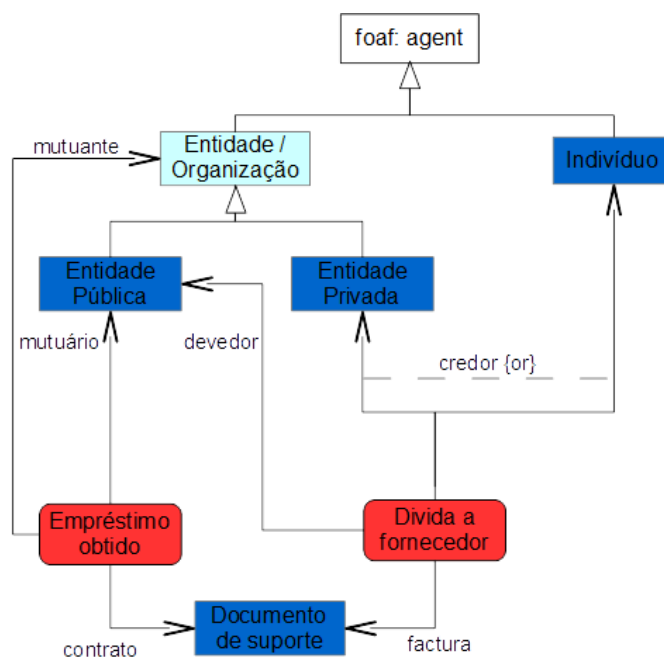


Figura 1 – Parte da ontologia proposta - *endividamento*

É de referir ainda que, para além de não serem apresentados atributos associados às entidades (tipos de dados) representadas, também não são apresentados atributos das relações (cardinalidade, obrigatoriedade) geralmente presentes em diagramas Entidade/Relacionamento. Assim, deste diagrama ressaltam dois tipos de ação (empréstimos (obtidos) e dívidas a fornecedores – representados por retângulos arredondados com a cor vermelha) bem como a sua relação (setas) com diferentes tipos de agentes e objetos (retângulos a azul), indicando o papel que representam em cada ação (texto associado às setas). O diagrama também mostra alguma hierarquia (relações de classe/subclasse representadas pelo triângulo) que pode existir entre elementos da ontologia. Procura-se ainda “enraizar” os elementos da ontologia proposta em outras ontologias de topo já existentes (retângulo sem cor), como é o caso da FOAF¹⁹ (que define o termo/classe *agent*²⁰) que pode dar origem às subclasses representadas), um princípio fundamental do desenvolvimento de ontologias.

A ontologia, no seu todo, contém atualmente cerca de 52 classes e cerca de 100 relacionamentos.

Do resultado do processo de desenvolvimento fazem também parte as definições associadas a cada um dos elementos gráficos do diagrama (agentes, ações, objetos, ...). Assim, exemplificando com o caso do endividamento, definiu-se empréstimo da seguinte forma:

¹⁹ *Friend of a Friend* (<http://xmlns.com/foaf/spec/>)

²⁰ http://xmlns.com/foaf/spec/#term_Agent

Empréstimo: contrato pelo qual se regula o acordo estabelecido entre o mutuante (entidade que concede o empréstimo) e o mutuário (entidade que o recebe) relativo a um financiamento, e onde se especificam todas as suas condições (montante, taxas de juro, ...) [adaptada da definição do Glossário do Banco de Portugal²¹]

A ontologia fica assim completa (com um formalismo semi-formal) com a definição de todos os conceitos associados.

5. CONCLUSÕES E TRABALHO FUTURO

Parece hoje incontestável uma aposta generalizada em formas de governação mais abertas aos cidadãos, nomeadamente no que diz respeito à possibilidade de estes poderem analisar e avaliar a atuação dos diferentes agentes públicos no que diz respeito à gestão dos recursos de que dispõem. Testemunho dessa maior vontade de transparência visando a *accountability* é o surgimento de inúmeros portais governamentais de disponibilização de dados.

Perante este cenário duas questões se colocam: Que itens de informação afinal disponibilizar? Como dar coerência global à fragmentação dos dados disponibilizados, quer através de diferentes portais (associados a diferentes temas), quer através de diferentes *datasets* num mesmo portal?

Neste contexto, a definição de uma ontologia de transparência que vise a *accountability* pode contribuir para dar resposta a estas duas questões tendo-se constatado que, apesar das diferentes ontologias relacionadas com governo eletrónico, nenhuma foi ainda desenvolvida com este objetivo específico. Assim, este trabalho visou concretamente colmatar esta lacuna.

O processo de desenvolvimento da ontologia não decorreu de forma puramente sequencial, já que em algumas fases foi necessário reformular/refinar o trabalho realizado em fases anteriores. Ao longo do desenvolvimento foram tomadas algumas opções que necessitarão de ser alvo de debate mais aprofundado, podendo inclusivamente, numa implementação futura, ser alvo de alteração. Um exemplo deste tipo de opções relaciona-se com a forma de tratar os dados relativos aos indivíduos visados em algumas das ações realizadas, nomeadamente o balanço entre o direito dos cidadãos à informação e a obrigatoriedade de prestação de contas pelos agentes públicos, por um lado, e o direito à privacidade desses indivíduos por outro (na publicitação, por exemplo, de processos disciplinares). Na análise deste aspeto pode inclusivamente ser necessário considerar questões legais que estão fora do âmbito deste trabalho. Também o nível de desagregação de alguns tipos de ações a reportar pode ser revisto, possivelmente no momento da definição dos atributos a disponibilizar.

²¹ <http://www.bportugal.pt/pt-PT/Glossarios/Paginas/Glossario.aspx>

No que diz respeito ao trabalho futuro, é de destacar a necessidade de proceder a uma validação externa da ontologia resultante. Assim, a ontologia desenvolvida não foi, ainda, sujeita a um processo formal de avaliação conforme requerido pela DSR, estando prevista a realização de uma avaliação externa através de entrevistas ou *focus groups* formados por jornalistas (considerando o seu papel de *information brokers* e, por isso, potenciais utilizadores diretos da ontologia desenvolvida) e especialistas em administração pública. A realização futura destas entrevistas ou *focus groups* não visa apenas a avaliação da ontologia agora proposta, mas também o seu desenvolvimento com a eventual integração de novas expressões/termos, identificação de outras relações entre eles, melhoria das definições que lhe estão associadas, bem como a identificação dos seus atributos. Neste momento o grau de formalismo da ontologia proposta é o considerado necessário para facilitar a sua avaliação e o seu refinamento por agentes externos ao processo de desenvolvimento. No futuro pretende-se aumentar o grau de formalismo para, pelo menos no caso português, possibilitar a utilização da ontologia enquanto suporte para implementar um sistema que, recorrendo a diferentes fontes de dados já existentes, disponibilize, de forma mais coerente, os dados considerados relevantes.

Em geral, o trabalho desenvolvido pretende contribuir para a melhoria da transparência na Administração Pública (e da sua avaliação), ao propor um conjunto de itens de informação que, ao centrar-se na perspectiva dos cidadãos, visa melhorar a *accountability* das entidades públicas. Em concreto, a ontologia aqui proposta poderá contribuir para uma melhor estruturação das diversas fontes de dados (portais) criadas ao abrigo dos atuais programas de governo aberto, possibilitando um mais fácil acesso, compreensão e análise dos dados por parte dos cidadãos ou de *information brokers* que atuem em sua representação.

Agradecimentos. Este trabalho foi parcialmente financiado pela Fundação para a Ciência e a Tecnologia (FCT) no âmbito do Projecto Estratégico PEst-OE/EEI/UI308/2014.

REFERÊNCIAS

- Bovens, M., "Public Accountability", in E. Ferlie, L. Lynne and C. Pollitt (Eds.), *The Oxford Handbook of Public Management*, Oxford University Press, Oxford, 2005, 182-208.
- Bovens, M., "Analysing and Assessing Accountability: A Conceptual Framework", *European Law Journal*, 13, 4 (2007), 447-468.
- Brusa, G., Caliusco, M.L. e Chiotti, O., *A process for building a domain ontology: an experience in developing a government budgetary ontology*, Proceedings of the second Australasian workshop on Advances in ontologies-Volume 72, 2006.

- Fung, A., "Infotopia: Unleashing the Democratic Power of Transparency", *Politics & Society*, 41, 2 (2013), 183-212.
- Gruber, T.R., "A translation approach to portable ontology specifications", *Knowledge acquisition*, 5, 2 (1993), 199-220.
- Heald, D., "Fiscal transparency: concepts, measurement and UK practice", *Public Administration*, 81, 4 (2003), 723-759.
- Heald, D., "Why is transparency about public expenditure so elusive?", *International Review of Administrative Sciences*, 78, 1 (2012), 30-49.
- Hevner, A.R., March, S.T., Park, J. e Ram, S., "Design science in information systems research", *MIS Quarterly*, 28, 1 (2004), 75-105.
- Hsieh, H.-F. e Shannon, S.E., "Three approaches to qualitative content analysis", *Qualitative health research*, 15, 9 (2005), 1277-1288.
- Klímek, J., Knap, T., Mynarz, J., Necaský, M. e Svátek, V., *LOD2 Deliverable 9a.1.1: Framework for Creating Linked Data in the Domain of Public Sector Contracts*, 2012.
- Linders, D. e Wilson, S.C., *What is Open Government? One Year after the Directive*, 12th Annual International Conference on Digital Government Research (Dg.o'11), College Park, MD, USA, 2011.
- Lourenço, R.P., "Data disclosure and transparency for accountability: A strategy and case analysis", *Information Polity*, 18, 3 (2013a), 243-260.
- Lourenço, R.P., *Transparência para accountability: uma estratégia e uma aplicação ao caso português*, 13ª Conferência da Associação Portuguesa de Sistemas de Informação, Évora, Portugal, 2013b.
- Noy, N.F. e McGuinness, D.L., *Ontology development 101: A guide to creating your first ontology*, Stanford medical informatics technical report SMI-2001-0880, 2001.
- Obama, B., *Transparency and open government. Memorandum for the heads of executive departments and agencies*, 2009.
- Orszag, P., *Open government directive. Memorandum for the Heads of Executive Departments and Agencies*, <http://www.whitehouse.gov/open/documents/open-government-directive>, 2009.
- Peffer, K., Tuunanen, T., Rothenberger, M.A. e Chatterjee, S., "A design science research methodology for information systems research", *Journal of Management Information Systems*, 24, 3 (2007), 45-77.

- Rutherford, B.A., "Developing a conceptual framework for central government financial reporting: intermediate users and indirect control", *Financial Accountability & Management*, 8, 4 (1992), 265–80.
- Studer, R., Benjamins, V.R. e Fensel, D., "Knowledge engineering: principles and methods", *Data & knowledge engineering*, 25, 1 (1998), 161-197.
- Uschold, M. e Gruninger, M., "Ontologies: principles, methods and applications", *The knowledge engineering review*, 11, 02 (1996), 93-136.
- Uschold, M. e Gruninger, M., "Ontologies and semantics for seamless connectivity", *ACM SIGMod Record*, 33, 4 (2004), 58-64.
- Vitvar, T. e Nazir, S., *D3.2 – SemanticGov Architecture v2.0: Providing Integrated Public Services to Citizens at the National and Pan-European level with the use of Emerging Semantic Web Technologies (SemanticGov)*, 2007.