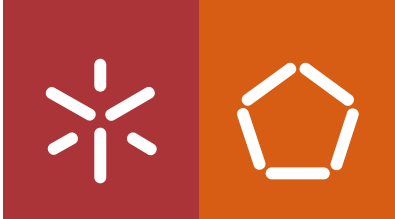


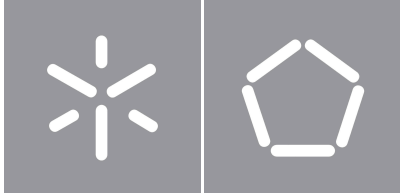


Quitéria Guimarães Pinheiro

Plataforma Colaborativa de Monitorização, em Tempo Real, de análises clínicas, nas Instituições de Saúde

Universidade do Minho
Escola de Engenharia





Universidade do Minho
Escola de Engenharia

Quitéria Guimarães Pinheiro

**Plataforma Colaborativa de Monitorização, em
Tempo Real, de análises clínicas, nas Instituições
de Saúde**

Dissertação de Mestrado
Mestrado em Bioinformática

Trabalho efetuado sob a orientação de
Professor Doutor José Manuel Ferreira Machado

Direitos de Autor e Condições de Utilização do Trabalho por Terceiros

Este é um trabalho académico que pode ser utilizado por terceiros desde que respeitadas as regras e boas práticas internacionalmente aceites, no que concerne aos direitos de autor e direitos conexos.

Assim, o presente trabalho pode ser utilizado nos termos previstos na licença abaixo indicada.

Caso o utilizador necessite de permissão para poder fazer um uso do trabalho em condições não previstas no licenciamento indicado, deverá contactar o autor, através do RepositóriUM da Universidade do Minho.

Licença concedida aos utilizadores deste trabalho:



CC BY-ND

<https://creativecommons.org/licenses/by-nd/4.0/> [Esta licença permite que outras pessoas usem o seu trabalho para qualquer fim, incluindo para fins comerciais. Contudo, o trabalho, na forma adaptada, não poderá ser partilhado com outras pessoas e têm que lhe ser atribuídos os devidos créditos.]

Agradecimentos

Após uma longa jornada de cinco anos, o presente projeto marca o encerramento de mais um ciclo. Ao longo destes anos, contei com diversos suportes, aos mais diversos níveis e, por isso, gostaria de deixar alguns agradecimentos.

Primeiramente, quero expressar a minha sincera gratidão ao meu orientador, Professor Doutor José Machado, pela sua constante presença e apoio ao longo deste ano. Agradeço imenso pelas oportunidades que me proporcionou desde o início, ao integrar a sua equipa, oportunidades essas que foram incrivelmente enriquecedoras e instrutivas.

Um agradecimento especial à Regina, que acompanhou de perto todas as etapas deste projeto e por encontrar sempre tempo para mim, mesmo diante das suas próprias responsabilidades. A tua ajuda foi fundamental para o sucesso deste projeto.

Aos meus pais e irmão, obrigada por estarem sempre lá para mim, por terem uma fé em mim que muitas vezes supera a minha, por confiarem sempre e me darem liberdade total para correr pelos meus sonhos, com a certeza de que será o caminho certo. Agradeço por nunca julgarem as minhas escolhas e, acima de tudo, por serem colo e casa, sempre! Vocês são o melhor lugar do mundo. Um 'obrigada' não chega para descrever o quão grata eu sou por vos ter na minha vida.

Um agradecimento especial ao meu namorado, por todo o respeito, apoio e companheirismo que partilhamos ao longo dos últimos dois anos. A tua presença foi fundamental em cada desafio que surgiu, no decorrer deste ano. Que venham os próximos desafios, que nós estamos prontos para os ultrapassar, juntos!

Aos meus amigos, àqueles que já vinham na bagagem antes de entrar na universidade e a todos aqueles que conheci por cá, ao longo deste percurso e que levo para a vida. Cada um de vocês, deixou uma marca única ao longo destes anos, enriquecendo-os com vivências e experiências que nunca irei

esquecer. Obrigada por fazerem parte da minha jornada e por tornarem cada momento tão especial.

Também quero expressar a minha gratidão à Universidade do Minho, que foi fundamental para o meu crescimento ao longo destes cinco anos. À Escola de Ciências, que foi a minha primeira casa e onde me licenciiei em Bioquímica e, em particular, à Escola de Engenharia, que se revelou, sem dúvida, o melhor caminho que eu poderia ter escolhido.

Declaração de Integridade

Declaro ter atuado com integridade na elaboração do presente trabalho académico e confirmo que não recorri à prática de plágio nem a qualquer forma de utilização indevida ou falsificação de informações ou resultados em nenhuma das etapas conducente à sua elaboração.

Mais declaro que conheço e que respeitei o Código de Conduta Ética da Universidade do Minho.

Universidade do Minho, Braga, janeiro 2024

Quitéria Guimarães Pinheiro

Resumo

Na era atual, marcada pela rápida evolução tecnológica, testemunhamos transformações profundas que permeiam todos os setores da sociedade. A incessante inovação tecnológica tem redefinido a forma como vivemos, trabalhamos e interagimos, proporcionando uma revolução que molda a essência da vida moderna.

No âmbito da saúde, essa revolução tecnológica não é apenas evidente, como também tem o potencial de redefinir padrões e expectativas. Os avanços tecnológicos na área da saúde oferecem benefícios significativos, desde diagnósticos mais precisos a tratamentos mais personalizados. A integração de tecnologias como inteligência artificial e *big data*, não apenas amplia o alcance dos serviços de saúde, como também melhora a eficiência e a qualidade do atendimento.

Os *softwares* desempenham um papel central nesta transformação, proporcionando ferramentas poderosas para profissionais de saúde. A gestão eficiente de dados, análises avançadas e aprimoramento da comunicação entre equipas médicas são apenas alguns dos benefícios oferecidos por estas soluções. A relevância crescente destes *softwares* destaca-se na contribuição das operações clínicas e na melhoria dos processos de tomada de decisão, resultando numa prestação de cuidados mais eficaz e centrada no paciente.

A elaboração da presente dissertação considerou o estudo minucioso da seguinte questão de investigação: "Qual o impacto da utilização de uma aplicação de análise de dados no domínio da medicina e na melhoria da qualidade de vida dos pacientes?" Para responder a esta questão, foi desenvolvida a *Analytic-Care*, uma aplicação de monitorização de análises clínicas, com a finalidade de ampliar a capacidade de comparação entre pacientes, proporcionando aos profissionais de saúde uma ferramenta valiosa. Este artefacto representa uma resposta às demandas crescentes por soluções que otimizem a análise comparativa de casos clínicos. Por fim, foi realizada uma **Prova de Conceito** - Análise **SWOT**, para avaliar a eficiência da aplicação, destacando não apenas a sua viabilidade técnica, mas também a sua capacidade de promover avanços na eficácia dos cuidados de saúde, contribuindo para um ambiente clínico mais sofisticado e integrado.

Palavras-chave instituições de saúde, monitorização, plataforma *web*, sistemas de informação em saúde, tomada de decisões médicas.

Abstract

In the current era marked by rapid technological evolution, we witness profound transformations that benefit all sectors of society. Relentless technological innovation has redefined how we live, work, and interact, bringing about a revolution that shapes the essence of modern life.

In the realm of healthcare, this technological revolution is not only evident but also has the potential to redefine standards and expectations. Technological advances in healthcare offer significant benefits, from more precise diagnostics to personalized treatments. The integration of technologies such as artificial intelligence and big data not only expands the reach of healthcare services but also improves efficiency and the quality of care.

Software plays a central role in this transformation, providing powerful tools for healthcare professionals. Efficient data management, advanced analytics, and improved communication among medical teams are just a few of the benefits offered by these solutions. The growing relevance of such software stands out in contributing to clinical operations and improving decision-making processes, resulting in more effective and patient-centered care.

The preparation of this dissertation considered the detailed study of the following research question: "What is the impact of using a data analysis application in the field of medicine and improving patients' quality of life?" To answer this question, AnalyticCare, a clinical analysis monitoring application, was developed with the aim of expanding the ability to compare patients, providing healthcare professionals with a valuable tool. This artefact represents a response to the growing demands for solutions that optimise the comparative analysis of clinical cases. Finally, a Proof of Concept - SWOT Analysis was carried out to assess the efficiency of the application, highlighting not only its technical feasibility, but also its ability to promote advances in healthcare efficiency, contributing to a more sophisticated and integrated clinical environment.

Keywords health information systems, health organizations, medical decision making, monitoring, web platform.

Conteúdo

1	Introdução	1
1.1	Contextualização e Motivação	1
1.2	Objetivos	2
1.3	Organização da Dissertação	3
2	Estado da Arte	5
2.1	Sistemas de Informação Hospitalar	5
2.2	Análise de <i>Big Data</i>	8
2.3	Tipos de repositórios de Base de Dados	9
2.3.1	Data Warehouse	9
2.3.2	Data Lake	10
2.3.3	Diferenças entre Data Warehouse e Data Lake	11
2.4	Processos ETL na área da saúde	11
2.5	Utilização de Ferramentas Colaborativas na Saúde	13
2.6	Plataformas de Visualização em Tempo Real	13
2.6.1	Tipos de técnicas de visualização de dados	14
2.6.2	Processos de visualização de dados de saúde	14
2.6.3	Painel de visualização interativo	15
2.6.4	Vantagens das técnicas de visualização na área da saúde	16
3	Metodologias de Investigação e Ferramentas de Desenvolvimento	19
3.1	Frameworks, bibliotecas e linguagens de programação	19
3.1.1	Bases de Dados	19
3.1.2	<i>Web Services</i> - Do início ao domínio do REST	21
3.1.3	Node.js	23
3.1.4	Bibliotecas	24

3.2	Metodologia da Prova de Conceito	25
4	Caso de Estudo I - Plataforma Colaborativa de Monitorização, em Tempo Real, de análises clínicas, nas Instituições de Saúde	27
4.1	Introdução e contextualização do problema	27
4.2	Análise de Requisitos	28
4.3	Origem do <i>dataset</i>	29
4.4	Arquitetura	31
4.4.1	Nome da aplicação: <i>AnalyticCare</i>	32
4.4.2	Página <i>Sign In</i> e página <i>Logout</i>	33
4.4.3	Página inicial	34
4.4.4	Página de detalhe do paciente	36
4.4.5	Tabela de exames do paciente	37
4.4.6	Página de tendência por idade	39
4.4.7	Página de tendência por localidade	41
4.5	Resultados e discussão	42
5	Prova de conceito	44
6	Conclusões e Trabalho Futuro	46
6.1	Principais contributos	47
6.2	Perspetiva para trabalho futuro	47

Lista de Figuras

1	Sistema de Informação: Integração entre recursos tecnológicos e funções.	8
2	Esquema simplificado de um processo Extract, transform, load (ETL)	12
3	Aplicações da base de dados SQLite. Adaptado de [1].	22
4	Esquema simplificado de uma arquitetura Rest.	23
5	Ilustração da Matriz representativa das principais características da análise SWOT	26
6	Arquitetura da aplicação <i>web</i> desenvolvida.	33
7	Logótipo da aplicação <i>web</i> desenvolvida.	33
8	Páginas iniciais com sistema de segurança, onde o médico efetua o seu registo na aplicação (esquerda) e, onde insere os seus dados sempre que acede à aplicação (direita). .	34
9	Página principal da plataforma que surge após registo na aplicação.	35
10	Página de detalhe do paciente.	36
11	Informações cruciais do paciente.	37
12	Quantidade de designações analisadas, ao longo dos exames realizados pelo paciente. .	38
13	Variação de todos os valores de certa designação, realizados em três exames distintos. .	40
14	Tendência por idade.	41
15	Tendência por localidade.	42

Acrónimos

APIs Interfaces de Programação de Aplicações.

BD Base de Dados.

COBRA Constraint-Based Reconstruction and Analysis.

COVID-19 Doença por Coronavírus 2019.

CRT Cálculo Relacional de Tuplas.

CSV Comma Separated Values.

DBMS sistemas de gestão de base de dados corporativos.

DL Data Lake.

DOM Document Object Model.

DSA Data Staging Area.

DW Data Warehouse.

EBI Instituto Europeu de Bioinformática.

ETL Extract, transform, load.

HBS Harvard Business School.

HTTP Hypertext Transfer Protocol.

IBM International Business Machines.

IS Instituições de Saúde.

IU Interface de Utilizador.

JASP Jeffreys's Amazing Statistics Program.

JSON JavaScript Object Notation.

LOINC Logical Observation Identifiers Names and Codes.

ML Machine Learning.

NCBI Centro Nacional de Informação Biotecnológica.

NPM Node Package Manager.

PoC Prova de Conceito.

REST Trasnferência Representacional de Estado.

SGBD Sistemas de gestão de base de dados.

SGBDR Sistemas de Gestão de Bases de Dados Relacionais.

SI Sistemas de Informação.

SIH Sistemas de Informação Hospitalar.

SOAP Simple Object Access Protocol.

SPSS Statistical Package for Social Science.

SQL Structured Query Language.

SRI Stanford Research Institute.

SSD Solid State Drive.

SWOT Strenghts, Weaknesses, Opportunities and Threats.

TI Tecnologias de Informação.

URL Uniform Resource Locator.

USB Universal Serial Bus.

WEB World Wide Web.

XML Extensible Markup Language.

Capítulo 1

Introdução

1.1 Contextualização e Motivação

Os sistemas de informação na saúde produzem e armazenam volumes significativos de dados eletrónicos. Os profissionais de saúde estão, diariamente, expostos a uma grande variedade de informações provenientes de várias fontes que necessitam de análise de forma a suportarem o constante processo de tomada de decisão [2]. Estas informações poderão incluir desde relatórios médicos eletrónicos e monitorização de pacientes até resultados laboratoriais, dados de ventiladores mecânicos e recomendações interprofissionais.

A capacidade de recuperar, sincronizar e partilhar dados é essencial tanto para a documentação clínica quanto para a solução de problemas, em tempo real, para pacientes individuais e para a população da unidade de terapia intensiva como um todo. Avanços tecnológicos que promovem a integração de dados relevantes numa única exibição ou “painel” podem melhorar a eficiência da equipa, acelerar decisões, agilizar processos de fluxo de trabalho e reduzir descuidos e erros na prática clínica. Estes dispositivos de monitorização de bem-estar e aplicações de *software* complementares capazes de gerar alertas e partilhar dados confiáveis de saúde já se encontram disponíveis [3].

Assim, tem sido notória a evolução que o setor da saúde tem registado ao longo das últimas décadas, sendo esta uma área em constante desenvolvimento em que a gestão da informação assume um papel fulcral nos **Sistemas de Informação (SI)**, instalados nas **Instituições de Saúde (IS)**. Entende-se por **SI**, sistemas automatizados, ou manuais, que devem ser capazes de recolher, processar, transmitir e difundir dados. Nas **IS** coexistem vários **SI**, com o intuito de facilitar as integrações do volume crescente de informação e, para tal estes sistemas utilizam tecnologias e normas diferentes. O desenvolvimento dos **SI** pretende promover a eficiência das organizações e é necessário que todos os intervenientes neste processo caminhem na mesma direção, melhorando a documentação, comunicação e coordenação entre sistemas.

1.2 Objetivos

O desenvolvimento desta dissertação tem como objetivo principal a resposta a uma questão de investigação. Esta questão foi colocada para compreender de que forma a utilização de uma aplicação de análise de dados na área da medicina tem impacto na melhoria da qualidade de vida dos utentes. Ao abordar este ponto, procura-se fornecer *insights* valiosos que possam informar práticas clínicas mais eficazes e promover avanços significativos no cuidado ao paciente.

Para isto, pretende-se proceder à análise e processamento de um *dataset*, inicialmente, com mais de 2 milhões de registos. Este *dataset* foi gerado de forma sintética, com o propósito de oferecer uma representação simulada e controlada de diversos contextos reais. Posteriormente, estes dados, serão consumidos pela plataforma elaborada, que irá monitorizar os resultados de análises clínicas, de cada paciente, e verificar de que forma os valores variam ao longo do tempo. A aplicação irá permitir também fazer comparações entre valores de análises de pacientes da mesma faixa etária, e comparações tendo em conta a localidade. Posto isto, o artefacto desenvolvido será capaz de ajudar os profissionais de saúde a tomar melhores decisões médicas, o que impactará positivamente o prognóstico e estado do paciente e garantirá um melhor desempenho nas organizações de saúde.

Com o propósito de alcançar esta finalidade, foram delineados os seguintes objetivos específicos, no início do processo de desenvolvimento da dissertação:

- Identificação e caracterização do problema;
- Pesquisa de Estado de Arte;
- Identificação da fonte de dados a utilizar;
- Estudo e caracterização dos dados provenientes;
- Pré-processamento de dados, com recurso a *Python*;
- Desenvolvimento dos modos da plataforma colaborativa para monitorização dos resultados de análises clínicas:
 - Vista orientada ao paciente;
 - Vista orientada à faixa etária;
 - Vista orientada à localidade;
- Elaboração da prova de conceito e/ou análise **SWOT** do artefacto desenvolvido.

1.3 Organização da Dissertação

A presente dissertação está estruturada de maneira a proporcionar uma compreensão lógica e detalhada do tema abordado e, encontra-se dividida em 6 capítulos.

Neste primeiro capítulo é apresentada a parte introdutória ao tema da dissertação, explicando uma breve contextualização e motivação. São também definidos os principais objetivos deste projeto.

O segundo capítulo corresponde ao estado da arte onde o tema deste estudo é aprofundado sendo constituído pela componente crítica das publicações científicas que fundamentam o trabalho. Neste capítulo são apresentados os principais conceitos e definições do mesmo e tópicos reagrupados segundo temas tratados, como por exemplo:

- Sistemas de Informação Hospitalar - Integração entre recursos tecnológicos e funções;
- Análise de *Big Data* e os seus desafios;
- Tipos de Repositórios - Aborda as principais diferenças entre *Data Warehouse* e *Data Lake*;
- Utilização de Ferramentas colaborativas nas instituições de saúde;
- Processos **Extract, transform, load (ETL)** - Apresenta a importância dos métodos de extração, transformação e carregamento dos dados bem como o momento em que este processo é executado;
- Plataformas de Visualização em Tempo Real - Expõe as vantagens da utilização destas plataformas no contexto da medicina, incluindo os tipos de técnicas de visualização de dados e os seus processos de visualização.

No terceiro capítulo, intitulado por 'Metodologias de Investigação e Ferramentas de Desenvolvimento', serão detalhadas as abordagens metodológicas utilizadas na condução da investigação, bem como as ferramentas de desenvolvimento empregues na criação da plataforma colaborativa de monitorização de análises clínicas.

O capítulo seguinte, quarto capítulo, referente ao 'Caso de Estudo I – Plataforma Colaborativa de Monitorização, em Tempo Real, de análises clínicas, nas Instituições de Saúde' apresentará uma análise detalhada do caso de estudo, destacando a arquitetura da plataforma, a sua explicação detalhada e os resultados obtidos.

No capítulo 5 - 'Prova de Conceito', será realizada uma demonstração prática da eficácia da plataforma desenvolvida, explorando casos específicos e validando os resultados obtidos.

Finalmente, o sexto capítulo, intitulado por 'Conclusões e Trabalho Futuro', proporcionará uma síntese das descobertas e contribuições do trabalho, juntamente com reflexões sobre possíveis direções futuras. Este capítulo consolidará o conhecimento adquirido ao longo da dissertação e delineará as oportunidades para investigação adicional.

Capítulo 2

Estado da Arte

Os profissionais de saúde, em ambientes hospitalares, devem coordenar todos os aspetos relacionados com o cuidado de um ou mais pacientes. Os dados clínicos são constantemente recuperados, documentados, analisados e comunicados a terceiros, todos inseridos na rotina diária de cuidados prestados pelos profissionais de saúde. Além disso, muitos dispositivos e monitores utilizados para auxiliar no tratamento de pacientes possuem sistemas de alarmes, que devem ser avaliados ao longo do dia de trabalho, juntamente com as ações que devem ser tomadas com base nas condições do paciente e em outros dados relevantes. A complexidade de tais processos de cuidado apresenta diversas oportunidades para possíveis descuidos em relação a detalhes importantes. A capacidade de vincular sistematicamente e logicamente diferentes conjuntos de dados num sistema de painel coeso é uma promessa significativa para melhorar a qualidade do atendimento, a segurança do paciente e os resultados clínicos num ambiente hospitalar [3].

2.1 Sistemas de Informação Hospitalar

Num âmbito global, as organizações de saúde estão empenhadas a aprimorar a qualidade do atendimento ao paciente, aumentar a satisfação do paciente e conter custos. Para atingir estes objetivos, estas organizações usam **Tecnologias de Informação (TI)** em saúde, incluindo **Sistemas de Informação Hospitalar (SIH)**, para gerir informações de pacientes - recolha, processamento, armazenamento e troca de informações - a alta velocidade [4, 5]. Um **SIH** é um sistema computacional projetado para gerir as atividades administrativas, onde se pretende registar dados demográficos dos doentes e dados do funcionamento da instituição; atividades financeiras, quando se registam dados relativos a custos ou receitas; verificação de *stock*, sendo importantes para fazer a gestão do mesmo e, atividades clínicas, quando se pretende registar os dados de saúde e doença dos utentes [6, 7, 8]. Assim, o sistema de informação é direcionado a pessoal administrativo, profissionais de saúde, gestores e/ou utentes.

De acordo com a finalidade a que se propõe, o **Sistemas de Informação** pode melhorar a qualidade do atendimento ao paciente, aumentando o acesso oportuno e a precisão das informações clínicas e administrativas relacionadas com o paciente. Este sistema aumenta a produtividade através do controlo de custos e melhora a qualidade do atendimento ao paciente, promovendo a troca de informações entre funcionários e departamentos do hospital. O uso de **Sistemas de Informação Hospitalar** também é eficaz na redução de erros médicos, facilitando o acesso adequado a informações de alta qualidade [9, 10, 11, 12, 13]. A área dos **Sistemas de Informação** está em constante crescimento, tanto nos cuidados primários como nos hospitais, e apesar de muito do esforço da informatização estar centrado nestes últimos, é aumentando a eficiência dos cuidados primários que se obtém um impacto positivo nos custos da prestação de cuidados de saúde [14].

Relativamente aos cuidados primários, sabe-se que os **Sistemas de Informação** clínicos auxiliam a gestão clínica do paciente no que diz respeito à prática médica, incluindo funcionalidades como registo de consultas e de tratamentos, dados laboratoriais e prescrição médica. Por outro lado, os **SI** administrativos suportam o funcionamento da instituição e gestão dos pacientes, incluindo funcionalidades como gestão de doentes, agendamento, orçamento e arquivo documental. Alguns sistemas incluem os **Sistemas de Informação** clínicos e administrativos num único sistema [14].

Estes sistemas representam as soluções informáticas mais comuns desenvolvidas para apoiar os serviços de saúde, abrangendo diversas atividades hospitalares em diferentes níveis de cuidados de saúde e tratamentos. Estão também responsáveis por armazenar informações clínicas e pessoais altamente sensíveis dos pacientes, sendo considerados recursos cruciais para os utilizadores destes **Sistemas de Informação** [15, 16]. No entanto, a natureza valiosa e sensível desses dados também implica riscos potenciais de violação da privacidade, confidencialidade e segurança das informações. Para mitigar estes riscos, os sistemas devem ser projetados para proteger contra acessos não autorizados, implementando robustos mecanismos de segurança, como métodos avançados de identificação e autenticação. É imperativo que tais sistemas permitam o acesso apenas a indivíduos autorizados, garantindo assim, a integridade e a confidencialidade dos dados e informações [17].

Sistemas de Informação: Integração entre recursos tecnológicos e funções

Como vimos anteriormente, um sistema de informação tem como finalidade transformar dados em informações qualificadas e contextualizadas, agregando significado aos dados. Este processo é resultado da interação entre cinco funções:

Entrada: Captura e armazenamento de dados;

Processamento: Conversão dos dados capturados em informação;

Armazenamento: Guarda os dados de forma organizada para utilização posterior;

Saída: Transfere informação produzida para o destino final;

Controlo do Desempenho: Analisa o funcionamento do sistema, através de relatórios emitidos na função de 'saída'. Se necessário, realiza ajustes nas funções de 'entrada' e 'processamento' para assegurar o alcance dos objetivos finais.

Com o desenvolvimento da informática, as funções dos **Sistemas de Informação** integram-se às plataformas de tecnologia de informação, compreendendo cinco tipos distintos de recursos:

Hardware: Constituídos por dispositivos físicos, utilizados no processamento de dados. Abrange processadores, disco rígido ou **SSD**, *pen drives* **USB**, impressoras, equipamentos médicos acoplados à rede de telecomunicações, entre outros;

Software: Conjunto de instruções de processamento de dados que engloba programas de computador. Estes programas dirigem e controlam as operações do *hardware*. O *software* também inclui procedimentos e documentações relacionadas ao processamento de dados, incorporando instruções operacionais e informações sobre a instalação e utilização do produto;

Recursos de dados: **Base de Dados** responsável por armazenar todos os dados do **Sistemas de Informação**;

Recursos de Rede: Refere-se ao meio físico de transmissão e de divulgação de dados que viabiliza a comunicação entre os diversos utilizadores do sistema;

Pessoas: Conferem sentido e significado aos **Sistemas de Informação**, na condição de utilizadores.

Ao integrarem-se com os recursos tecnológicos e funcionais, os **Sistemas de Informação** tornam-se elementos essenciais na estruturação e operação eficiente das organizações, promovendo a transformação de dados em informações valiosas para a tomada de decisão e o alcance dos objetivos estabelecidos [18]. A figura 1 ilustra a união entre as capacidades tecnológicas e as operações dos **SI**.

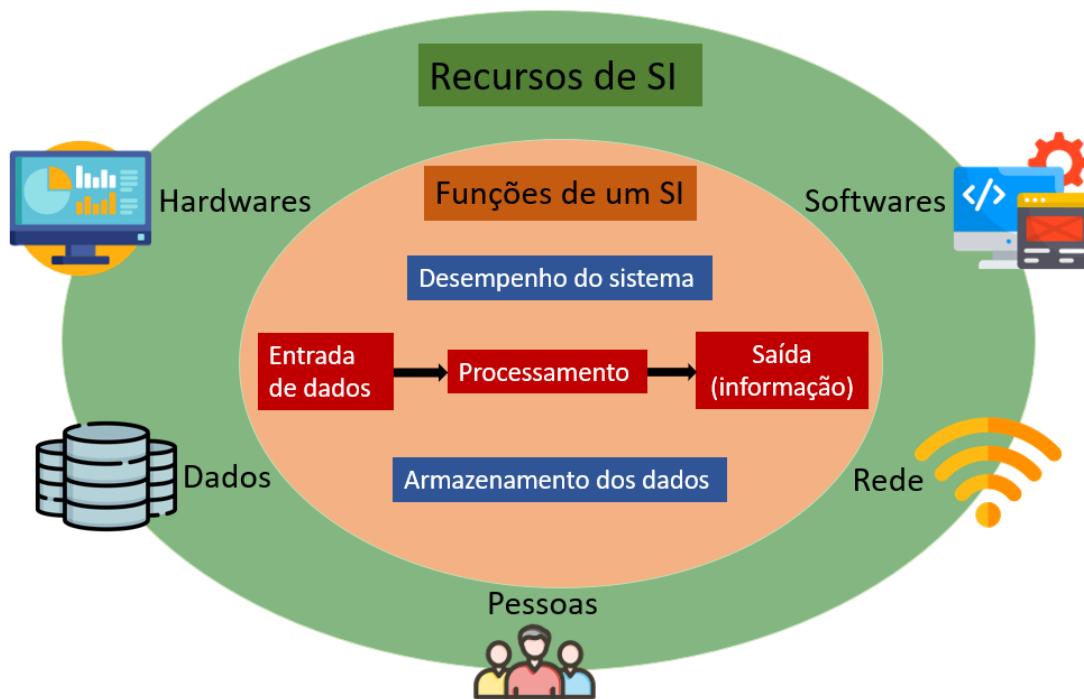


Figura 1: Sistema de Informação: Integração entre recursos tecnológicos e funções.

2.2 Análise de *Big Data*

Nos últimos anos, com o aparecimento de vários novos serviços de Internet, aplicações *web* e plataformas de comunicação social (por exemplo, *Facebook*, *Twitter*, *Instagram*, uma nova categoria de dados emergiu: o *Big Data*. Esta evolução apresenta novos desafios, especialmente em relação ao grande "volume" de dados, bem como à heterogeneidade destes dados, que podem ser estruturados, semi-estruturados e não estruturados (ou "variedade". Além disso, são impostos desafios quanto à "velocidade" necessária para processar esses dados e à confiabilidade dos dados (ou "veracidade". Estes termos representam os populares "4Vs" de *Big Data* na literatura [19]. Estes conjuntos de dados ultrapassam a capacidade das ferramentas tradicionais de capturar, armazenar, processar e analisar dados. Como resposta a este desafio, uma nova estratégia de armazenamento de dados foi proposta, conhecido como o processo **ETL** [20, 21]. Neste contexto, os dados são, inicialmente, extraídos das suas fontes (1. extrair) e armazenados no seu estado bruto num **Data Lake** (2. carregar), aplicando apenas pequenas transformações (3. transformar). Posteriormente, esses dados passarão por tarefas de transformação necessárias para serem finalmente carregados num **Data Warehouse** de destino, conforme exploraremos mais adiante.

2.3 Tipos de repositórios de Base de Dados

Os dados representam o núcleo essencial de qualquer organização nos dias de hoje. Reconhecendo a importância fundamental dos dados nos sistemas modernos de inteligência de negócios, as organizações procuram tomar decisões significativas e manter a competitividade no mercado. A eficiente e ideal análise de dados confere uma vantagem competitiva, o que influencia diretamente o desempenho e os serviços oferecidos. As principais organizações atualmente geram, recolhem e processam volumes substanciais de dados, que se enquadram na categoria de *Big Data*. A gestão e análise eficazes desse amplo espectro de *Big Data* representam um desafio complexo.

A utilização apropriada da vasta coleção de informações de uma organização pode fornecer *insights* valiosos sobre as estratégias de negócios. Nesse contexto, dois sistemas de gestão de dados proeminentes na análise de *Big Data* - *Data Warehouse* e *Data Lake* - atuam como plataformas para acumular os grandes volumes de dados gerados e utilizados pelas organizações. Embora possam parecer semelhantes à primeira vista, estes sistemas diferem em termos de características e aplicações [22].

2.3.1 *Data Warehouse*

O conceito de *Data Warehouse* foi pioneiramente introduzido no final da década de 1980 pelos pesquisadores da *International Business Machines* (IBM), Barry Devlin e Paul Murphy com o propósito de criar uma arquitetura capaz de resolver o fluxo de dados para ambientes de suporte à decisão [22]. Segundo a definição de Inmon, “um **DW** é uma coleção de dados orientada por assunto, não volátil, integrada e variável no tempo para apoiar as decisões de gestão” [23].

Formalmente, o *Data Warehouse* representa um amplo repositório de dados, permitindo o armazenamento e integração de informações provenientes de diversas fontes, de maneira bem estruturada. Este desempenha um papel crucial no processo de tomada de decisões, proporcionando uma análise de dados adequada [24].

No contexto da gestão de dados corporativos, o armazenamento de dados refere-se a um conjunto de sistemas de tomada de decisão projetados para capacitar especialistas em informações, como líderes, administradores ou analistas, visando melhorar e otimizar o tempo dedicado a tomar decisões estratégicas. Os sistemas de *Data Warehouse* atuam como ferramentas essenciais de inteligência de negócios, sendo amplamente adotados na gestão de dados, por médias e grandes empresas [25, 26].

Na última década observou-se um significativo desenvolvimento, tanto na diversificação de produtos e serviços oferecidos quanto na adoção em larga escala desses avanços pela indústria [22].

Na estrutura do **Data Warehouse**, os dados são regularmente extraídos de programas que suportam as operações comerciais e são consolidados em unidades de processamento especializadas. O **DW** torna-se, então, uma fonte primária de dados para a produção, análise e apresentação de relatórios, utilizando ferramentas como relatórios instantâneos, portais eletrônicos e leituras digitais. Os **DW**, por outro lado, estão intrinsecamente ligados à assistência à decisão, integrando dados provenientes de várias fontes e desempenham um papel crucial na análise, mineração de dados e origem de relatórios [22].

2.3.2 **Data Lake**

No início do século XXI, o advento de novos tipos de dados, em volumes crescentes provenientes da Internet e de interfaces com diversas empresas, como transações comerciais baseadas na *Web*, *streaming* em tempo real, dados de sensores e comunicação social, trouxe a necessidade de soluções mais eficientes para armazenar e analisar grandes quantidades de dados semi-estruturados e não estruturados. As abordagens tradicionais de “*schema-on-write*”, que envolve a criação de um esquema para dados antes de serem gravados na base de dados, mostraram-se ineficientes para gerir esta vasta quantidade de dados, levando ao surgimento de uma abordagem moderna conhecida como **Data Lake (DL)** [27, 28, 29].

Os **DL** são repositórios centralizados de armazenamento que permitem aos utilizadores armazenar dados brutos e não processados no seu formato original, incluindo dados não estruturados, semi-estruturados ou estruturados, em grande escala. Isto permite que as empresas façam escolhas de negócios mais esclarecidas como por exemplo, ao utilizar visualizações, painéis de análise de *Big Data*, **Machine Learning** e análise em tempo real [22].

Segundo Dixon, “enquanto um **DW** pode ser comparado a uma garrafa de água limpa e pronta para consumo, o **DL** é visto como um lago completo de dados num estado mais natural” [30]. Um **Data Lake** armazena dados na sua forma bruta original, variando em tamanho e estrutura, e carece de qualquer organização específica. Embora forneça flexibilidade e escalabilidade, a implementação e análise de dados nos **DL** também apresentam desafios [22].

Cada vez mais populares, os **Data Lake** oferecem às organizações uma forma centralizada de armazenar os seus dados. Podem conter dados não estruturados ou estruturados, permitindo que as organizações armazenem informações de diferentes fontes sem a necessidade de transformação prévia dos dados [27]. Os **DL** substituem a abordagem anterior de organizar e processar dados de várias fontes por um repositório centralizado, eficiente e flexível, possibilitando que as organizações maximizem ganhos e desenvolvam estratégias superiores, por exemplo, em áreas como *marketing*. Além disso, como já foi referido, os **DL** possibilitam escalabilidade, eliminando a necessidade de transformação complexa e pré-

processamento de dados associados aos **DW**. Desta forma, a sobrecarga financeira inicial da ingestão de dados é reduzida e após o armazenamento no **Data Lake**, os dados ficam disponíveis para análise por parte da organização.

2.3.3 Diferenças entre **Data Warehouse** e **Data Lake**

Embora **Data Warehouse** e **Data Lake** muitas vezes sejam usados como dois termos permutáveis, é importante destacar que eles não são sinônimos [22]. Uma das suas principais diferenças reside nas diferentes estruturas. Como mencionado anteriormente, um **DW** armazena dados processados e filtrados, enquanto que os **Data Lake** mantêm dados brutos ou não processados. Nos **Data Warehouse**, os dados são processados e organizados num único esquema antes de serem armazenados, enquanto que nos **Data Lake**, os dados brutos e não estruturados são inseridos e organizados conforme necessário.

Os **Data Lake** têm uma capacidade comparativamente maior do que os **Data Warehouse** e são ideais para a análise de dados brutos e não processados, bem como para a implementação de **Machine Learning**. Outra diferença fundamental reside no propósito ou objetivo de uso. Geralmente, os dados processados armazenados em **DW** são utilizados para finalidades específicas, otimizando o espaço de armazenamento. Por outro lado, o propósito de uso de **Data Lake** não é predefinido e pode servir a qualquer finalidade. Enquanto os **DW** podem ser acessados e utilizados por qualquer pessoa ou empresa sem a necessidade de conhecimento especializado, a análise de **DL**, por envolver dados não processados, requer a experiência de cientistas de dados ou ferramentas específicas para aplicações comerciais. A acessibilidade e facilidade de uso são aspetos adicionais que diferenciam **DW** e **DL**. A arquitetura de um **Data Lake** não possui uma estrutura rigidamente definida, oferecendo flexibilidade, enquanto que a estrutura de um **Data Warehouse** garante uma maior segurança, sendo mais resistente a manipulações externas [22].

2.4 Processos **ETL** na área da saúde

A globalização e a disseminação das **Tecnologias de Informação**, aliadas à necessidade de acesso rápido e fácil a informações confiáveis, levaram a líderes empresariais a substituir os sistemas tradicionais de computação empresarial por soluções mais avançadas de suporte à decisão [31]. A integração de dados em repositórios é conduzida por meio do processo de três fases conhecido como: extrair, transformar e carregar (**Extract, transform, load (ETL)**, em inglês). Este procedimento envolve a extração de dados de diferentes fontes, a sua transformação (abrangendo preparação, conversão, limpeza, filtragem,

junção, agregação, entre outros) e, por fim, o carregamento para um repositório de dados. A estrutura geral dos processos **ETL** compreende três etapas sequenciais - extrair, transformar e carregar - e três camadas principais - fontes de dados, área de preparação de dados e repositório de dados [32, 23]. A figura 2 mostra, de forma simplificada, como funcionam os processos **ETL**.

Na fase inicial de extração, os dados são retirados de diversas fontes de dados e convertidos para um formato adequado, facilitando o processamento durante a etapa de transformação. Estas fontes podem ser internas, como **sistemas de gestão de base de dados corporativos (DBMS)**, arquivos **XML**, arquivos simples (texto ou **CSV**), ou externas, como aplicações *Web*, sensores, câmaras, redes sociais, e-mails, entre outros. De acordo com [33], existem dois métodos físicos (extração *online* e *offline*) e dois métodos lógicos (extração completa e incremental). A extração completa, também conhecida como “extração inicial” [34], implica o primeiro carregamento da base de dados com dados de fontes operacionais. Já na extração incremental, conhecida como captura de dados alterados, atualiza a base de dados com dados modificados e sistemas de origem desde a última extração, seguindo um ciclo periódico alinhado às necessidades do negócio [34].

As transformações, denominadas “tarefas comuns” em [35, 36], ou “atividades”, conforme outros pesquisadores [37], incluem agregação, junção, conversão e filtragem. Estas operações são executadas na **Data Staging Area (DSA)** - a área de preparação e limpeza de dados - também conhecida como armazenamento intermediário de dados, antes de serem carregados na área de armazenamento de destino.

A última fase do processo **ETL** é o carregamento, que transfere os dados transformados para os repositórios de dados de destino [31]. Os dados, agora no formato final desejado, são carregados no repositório de dados de acordo com um modelo padrão que deve atender com precisão às necessidades iniciais do cliente [31].



Figura 2: Esquema simplificado de um processo **ETL**.

2.5 Utilização de Ferramentas Colaborativas na Saúde

As ferramentas colaborativas desempenham um papel cada vez mais essencial na era da informação, proporcionando uma plataforma dinâmica para a comunicação eficiente e a colaboração interdisciplinar. Além de facilitar a interação entre grupos, independentemente da localização geográfica, estas soluções promovem a integração de conhecimentos e a troca de ideias de forma rápida e acessível. A utilização eficiente destas ferramentas não apenas aprimora a disseminação de informações, como também impulsiona o trabalho intelectual ao proporcionar um ambiente propício à inovação e à criação colaborativa [38].

No contexto social das **Tecnologias de Informação**, observamos um impacto significativo em vários domínios. A pesquisa e partilha de informações foram transformadas, possibilitando a rápida propagação de descobertas e conhecimentos. A confiabilidade da informação é agora uma preocupação central, impulsionada pela necessidade de discernir entre fontes precisas e questionáveis. A melhoria na comunicação remota proporciona uma conectividade sem precedentes, permitindo colaborações globais e a coordenação de equipas distribuídas geograficamente [39].

Além disso, as **TI** desempenham um papel crucial nas novas dimensões da educação. A aprendizagem colaborativa é facilitada por plataformas *online*, proporcionando aos alunos e educadores oportunidades enriquecedoras de interação e partilha de conhecimentos. O acesso a recursos educacionais é expandido, promovendo a democratização do conhecimento [40].

Assim, as ferramentas colaborativas não apenas redefinem a forma como trabalhamos e comunicamos, mas também moldam profundamente a dinâmica social e educacional na era digital.

2.6 Plataformas de Visualização em Tempo Real

A visualização de dados no setor da saúde beneficia de tecnologias modernas sofisticadas, permitindo que profissionais de diversas áreas apresentem o seu trabalho e comuniquem informações de forma eficaz. As técnicas de visualização desempenham um papel crucial ao auxiliar os profissionais de saúde na compreensão das tendências passadas e presentes, assim como na previsão e antecipação de direções futuras. Geralmente, a visualização de dados envolve a representação de informações por meio de gráficos, tabelas, diagramas e imagens [41]. Tais técnicas proporcionam aos profissionais de saúde uma maneira intuitiva de identificar e entender tendências, valores discrepantes e padrões de dados [42].

A utilização efetiva de técnicas de visualização tem sido fundamental em vários setores da saúde,

desempenhando um papel fundamental no apoio à tomada de decisões clínicas importantes relacionadas à saúde do paciente e da comunidade. Ao sintetizar dados brutos em gráficos, as organizações de saúde podem facilitar a interpretação imediata de tendências e padrões [43].

2.6.1 Tipos de técnicas de visualização de dados

Segundo o estudo de Narayan et al. em 2021 [44], a acumulação de um grande volume de dados de saúde torna o conceito de *big data* muito comum nos setores de saúde, considerando os aspetos de volume, velocidade, variedade e veracidade dos dados [44], conforme analisado anteriormente.

As técnicas de visualização desempenham um papel crucial ao simplificar o *big data* e ao torná-lo mais fácil de interpretar, mesmo para profissionais que não são da área da saúde. É imperativo que os setores de saúde possuam ferramentas adequadas para o desenvolvimento de visualizações. Muitas ferramentas fazem uso de técnicas de visualização de dados, como tabelas dinâmicas e gráficos. Além disso, diferentes ferramentas de visualização estão disponíveis, atendendo tanto a utilizadores técnicos quanto não técnicos. Por exemplo, o *Microsoft Excel* fornece uma excelente ferramenta de visualização para profissionais de saúde com pouco ou nenhum conhecimento técnico.

Outro exemplo são *softwares* estatísticos como *International Business Machines (IBM) Statistical Package for Social Science (SPSS)* e *Jeffreys's Amazing Statistics Program (JASP)*. Estes *softwares* estatísticos fornecem sistemas de visualização interativos destinados a utilizadores especializados. Estas ferramentas de visualização desempenham um papel crucial na vigilância pública de doenças [44]. Para uma visualização eficiente e rápida de dados de saúde, é essencial considerar a divulgação das informações para as partes interessadas por meio de várias plataformas. Na saúde pública da comunidade, a partilha de dados e integração de informações visuais numa plataforma universal são benéficos e necessários. Diversas plataformas estão disponíveis para este fim, incluindo plataformas *cloud-based*. Além disso, estas plataformas, como por exemplo, *ParaView* e *Gephi*, permitem que os utilizadores criem e armazenem gráficos, facilitando a reprodutibilidade científica ao combinar as figuras científicas com os seus dados subjacentes e promovendo a discussão entre os colaboradores [44].

2.6.2 Processos de visualização de dados de saúde

A análise de dados proporciona aos profissionais de saúde uma visão abrangente e detalhada do cenário de saúde, permitindo uma compreensão mais profunda das dinâmicas e complexidades envolvidas. Além disso, a aplicação de técnicas de inteligência artificial amplifica a capacidade de identificar padrões não óbvios e correlações complexas nos dados de saúde [45].

A comparação contínua dos dados históricos não apenas destaca as mudanças ao longo do tempo, mas também fortalece a capacidade preditiva das análises. A identificação precoce de tendências emergentes é crucial para a implementação de intervenções proativas e estratégias de saúde pública eficazes. A visualização destas tendências em diferentes dimensões, como faixa etária, localidade e prevalência de doenças, proporciona *insights* valiosos para a adaptação contínua das práticas de saúde e otimização dos recursos [46].

Além disso, a interoperabilidade destas ferramentas de visualização com outros sistemas e registros de saúde contribui para a integração eficiente de dados provenientes de diversas fontes. Isto não só enriquece a análise, como também promove uma abordagem holística na compreensão dos padrões de saúde, considerando fatores multifacetados que impactam a prevalência e o tratamento de doenças. Esta integração harmoniosa de dados e análise fortalece a base de conhecimento para decisões informadas e estratégias de saúde mais eficazes [47].

2.6.3 Painel de visualização interativo

Conforme destacado por [48], os *dashboards* e as ferramentas de análise de dados, comumente integrados nos *softwares* existentes nos sistemas de saúde, desempenham um papel crucial ao combinar diversos relatórios interativos. Estes painéis são categorizados em três tipos principais: o tipo ativo, que apresenta dados, em tempo real, nas organizações de saúde; o tipo estratégico, que evidencia tendências ao longo do tempo; e o tipo analítico, que oferece análises mais avançadas.

Os *dashboards* ativos fornecem uma visão imediata em tempo real dos dados cruciais, como o fluxo de pacientes, o estado de procedimentos, entre outros. Esta funcionalidade permite uma resposta ágil a eventos que estão a decorrer, otimizando a eficiência operacional [49].

Por outro lado, os *dashboards* estratégicos são projetados para destacar tendências ao longo do tempo, oferecendo uma visão mais ampla e histórica dos dados. Estes painéis facilitam a identificação de padrões, variações sazonais e mudanças significativas nas métricas de saúde, sendo valiosos para o planeamento a longo prazo e a adaptação a mudanças no cenário da saúde [50].

Os *dashboards* analíticos, por sua vez, têm o propósito de fornecer análises avançadas, incorporando ferramentas sofisticadas de visualização e técnicas estatísticas para explorar relações complexas nos dados de saúde. Esta categoria permite análises aprofundadas, identificação de correlações entre diferentes variáveis e obtenção de *insights* detalhados para fundamentar decisões estratégicas e aprimorar a qualidade dos cuidados prestados. Estes *dashboards* são essenciais para a tomada de decisões informadas com base na compreensão mais profunda dos dados de saúde [51].

2.6.4 Vantagens das técnicas de visualização na área da saúde

A implementação e a adoção de diferentes técnicas de visualização nos setores da saúde são fundamentais para aprimorar a prestação geral de cuidados de saúde pelos profissionais do setor. Diversas instituições de saúde incorporaram ferramentas de análise de dados, incluindo **ML** e inteligência artificial. Os benefícios associados à implementação e adoção destas técnicas, em estabelecimentos de saúde, são diversos:

1. **Melhoria geral do atendimento ao paciente:** A utilização da visualização de dados de saúde tem impactos positivos no fornecimento global de cuidados de saúde. Desempenhando um papel crucial no apoio às tomadas de decisões clínicas dos profissionais de saúde, facilita a antecipação de ameaças e permite respostas imediatas. A identificação de várias medidas críticas possibilita a análise aprofundada da situação dentro de uma organização de saúde. A visualização, em tempo real, dos dados de saúde do paciente é essencial para aprimorar a qualidade do atendimento, permitindo que os profissionais tomem decisões clínicas com base na situação atual do paciente [52].

Vamos analisar um exemplo sucinto, que descreve perfeitamente a utilidade de técnicas de visualização: estas técnicas são fundamentais para monitorizar parâmetros de saúde e foram especialmente relevantes durante a pandemia da **COVID-19**. Um indicativo crucial a analisar pelos profissionais de saúde, durante esse período, foi a saturação de oxigénio. [53]. A visualização, em tempo real, desses dados foi essencial para avaliar as respostas dos pacientes ao tratamento, permitindo a monitorização dos níveis de saturação de oxigénio e ajustes terapêuticos conforme necessário [54].

Este exemplo ilustra a importância das aplicações, bem-sucedidas, de plataformas de visualização em tempo real, destacando a possibilidade de adaptar esta metodologia para análises clínicas específicas. Em diversas unidades de internamento, a visualização em tempo real é utilizada para monitorizar parâmetros como pulso, frequência cardíaca e a pressão arterial [55], contribuindo para a deteção precoce de parâmetros anormais e facilitando intervenções necessárias para melhorar resultados e a qualidade geral dos cuidados de saúde.

2. **Reconhecimento de tendências e padrões de doenças:** Identificar tendências e padrões representa outra vantagem significativa da utilização de técnicas de visualização nos setores de saúde. Esta capacidade é crucial para a tomada de decisões na prestação de cuidados de saúde, permitindo a deteção precoce de possíveis problemas e a implementação de estratégias preventivas. Um

aspecto crítico é a identificação e avaliação de padrões de doenças em populações específicas. O reconhecimento desses padrões fornece *insights* valiosos, indicando a necessidade de investigar os fatores subjacentes a uma elevação nas tendências. A análise de dados e a representação visual de indicadores como a obesidade exemplificam a identificação de padrões, desempenhando um papel crucial na saúde pública e comunitária. A compreensão dos fatores por trás dessas tendências pode promover a consciencialização sobre a necessidade de modificar estilos de vida, tanto a nível individual como coletivo [56].

3. **Apresentação de dados para diferentes audiências:** Frequentemente, a interpretação de dados de saúde é bastante desafiadora para profissionais de diversas disciplinas, principalmente para indivíduos sem formação médica. A principal finalidade da apresentação de dados é simplificar informações complexas, tornando-as compreensíveis para qualquer público, independentemente do seu conhecimento prévio. Por exemplo, ao exibir dados de prevalência de doenças em gráficos para o público em geral, ocorre uma simplificação eficaz que facilita a compreensão [41]. Além disso, a utilização de apresentações visuais para ilustrar fatores que impactam a prevalência de doenças tornam as informações médicas valiosas e facilmente acessíveis a qualquer público-alvo.

4. **Desempenho acelerado:** A visualização de dados, em tempo real, em organizações de saúde proporciona um desempenho mais rápido e eficiente nos cuidados de saúde prestados. Isto reflete-se na agilidade da tomada de decisões clínicas, especialmente em situações críticas, resultando em impactos positivos no prognóstico e estado de saúde do paciente.

Além disso, a atuação acelerada contribui para a redução de inadequações da assistência prestada, garantindo um melhor desempenho global das organizações de saúde. Este rápido e eficiente atendimento não apenas beneficia os pacientes, mas também constrói uma reputação sólida de eficácia para as instituições de saúde [42].

5. **Deteção de erros e fraudes:** Um dos benefícios mais proeminentes da implementação de diversas técnicas de análise e visualização de dados nos setores de saúde é capacidade de detetar fraudes e erros, especialmente aqueles relacionados à faturação médica. Conforme relatado pela *Medicare* e *Medicaid*, a maioria dos casos de fraude que ocorrem nas unidades de saúde são cometidas por profissionais de saúde internos, resultando em perdas financeiras substanciais para as instituições [56].

As formas proeminentes de fraude incluem cobranças duplicadas, cobranças fictícias, prescrições falsas e outros tipos de fraudes no seguro de saúde cometidas por prestadores de serviços. Uma

correlação efetiva entre as partes interessadas, como pacientes, pagadores de assistência médica (entidades ou organizações responsáveis por pagar ou cobrir custos dos serviços de assistência médica) e provedores de sinistros (organizações ou entidades que processam e geram reclamações apresentadas pelos beneficiários de serviços de saúde), pode melhorar a integridade do processo de cobrança e reduzir esquemas fraudulentos [57]. A implementação de técnicas de visualização de dados tem sido crucial para aumentar a transparência no setor da saúde.

Capítulo 3

Metodologias de Investigação e Ferramentas de Desenvolvimento

As indagações de pesquisa apresentadas nos capítulos anteriores conduziram à exploração de resoluções, bem como de abordagens metodológicas que contribuíssem para a criação de um conteúdo sólido e proveitoso. É evidente que este processo demanda uma pesquisa cuidadosa e uma análise detalhada das metodologias e tecnologias disponíveis, a fim de selecionar aquelas mais apropriadas para o desenvolvimento das soluções propostas. Nesse sentido, este capítulo procura apresentar uma descrição abrangente das metodologias e tecnologias aplicadas ao longo da condução deste estudo. É relevante observar que a seleção dos métodos e ferramentas não se fundamentam apenas nas vantagens da sua implementação, mas também na sua conformidade com os sistemas relacionados e nas respectivas limitações.

3.1 Frameworks, bibliotecas e linguagens de programação

3.1.1 Bases de Dados

Desde o início da revolução digital, o armazenamento de informações para consultas subsequentes tem sido um desafio persistente. Na indústria são gerados vários tipos de dados, tais como, dados semiestruturados e não estruturados. Mesmo que possa parecer simples à primeira vista, as bases de dados carregam uma grande responsabilidade, uma vez que o êxito das operações às quais são submetidas pode influenciar significativamente a confiabilidade e eficiência de um sistema como um todo.

Uma base de dados é uma estrutura organizada para armazenar e gerir dados de forma eficiente. Os dados podem ser inseridos, recuperados, atualizados e excluídos conforme necessário. [58]. Os **Sistemas de gestão de base de dados (SGBD)** são *softwares* projetados para criar e gerir estas bases de dados, oferecendo recursos como produção de relatórios e controlo de operações de leitura e gravação.

Atuam também como intermediários entre os utilizadores finais e a base de dados, facilitando a organização e manipulação das informações. As principais funções dos **SGBD** incluem a gestão dos dados, a definição do esquema da base de dados, que estabelece a sua estrutura lógica, e o mecanismo da base de dados, que permite o acesso, atualização e bloqueio dos dados. Estes três elementos fundamentais garantem a execução de procedimentos administrativos padronizados, garantindo simultaneidade, recuperação, segurança e integridade dos dados [59].

Os **Sistemas de Gestão de Bases de Dados Relacionais (SGBDR)** são uma versão avançada dos **SGBD**. Baseiam-se no modelo Relacional - álgebra relacional e **Cálculo Relacional de Tuplas (CRT)** - introduzido por Edgar Codd [60]. As bases de dados relacionais são altamente estruturadas em tabelas, ou seja, todos os dados da tabela são armazenados em linhas e colunas, frequentemente referidas como tuplas e atributos, respetivamente. Cada linha contém um valor de dados para a coluna correspondente, enquanto que cada coluna possui um tipo de dados normalizado, muitas vezes sujeitos a restrições [61].

As tabelas estão relacionadas entre si através de chaves primárias e estrangeiras garantindo a “integridade referencial” entre os objetos. Estas tabelas relacionais são a base do modelo relacional [62] e são amplamente utilizadas em várias bases de dados populares, atualmente.

Entre as características proeminentes dos sistemas **SGBDR**, destacam-se as seguintes:

- Armazenamento dos dados em tabelas, organizadas em colunas e linhas, fornecendo uma estrutura eficiente e organizada;
- Fornece diversas facilidades, como chave primária, que garante a identificação única das linhas, e utiliza uma coluna comum que pode ser partilhada entre mais do que uma tabela;
- Possibilidade de criação de índices para acelerar a recuperação de dados, aumentando a eficiência do sistema;
- Criação de tabelas virtuais, que simplificam as consultas ao longo do tempo, e oferecem acessibilidade multi-cliente controlada por utilizadores individuais;

Estes recursos tornam os **Sistemas de Gestão de Bases de Dados Relacionais** amplamente utilizados em todo o mundo.

Structured Query Language

O **SQL**, abreviação de *Structured Query Language*, refere-se a uma linguagem de programação utilizada para gerir dados em **Sistemas de Gestão de Bases de Dados Relacionais (SGBDR)**. Desenvolvida na década de 70, o **SQL** tornou-se a ferramenta padrão para administradores e programadores de base de dados [63]. Inclui a capacidade de inserir, consultar, atualizar e excluir dados, bem como gerir o esquema de dados (estruturas de base de dados) e controlar o acesso aos dados na base de dados. As operações e consultas em **SQL**, geralmente, estão sob formas de comandos, escritas como instruções [64].

A versatilidade do **SQL** tornou-o uma ferramenta essencial não apenas para consultas básicas, como também para tarefas avançadas de gestão e integração de dados. Com capacidades que vão desde consultas analíticas até manipulação de esquemas de dados, o **SQL** desempenha um papel fundamental em praticamente todas as etapas do ciclo de vida de uma base de dados [65]. De seguida, será descrita uma base de dados atual que se baseia num modelo relacional e será utilizada nesta dissertação.

Base de Dados SQLite

O SQLite é um mecanismo de base de dados incorporado conhecido pela sua eficiência e ampla gama de recursos. A sua conceção compacta e eficiente, torna-o uma escolha popular numa variedade de aplicações, incluindo dispositivos móveis [66, 67]. O *Android* está entre as plataformas mais conhecidas com alta dependência do SQLite, e os navegadores *Chrome* e *Firefox* também fazem uso do SQLite para manter objetos da *web* em armazenamento local [67]. Além disso, dispositivos médicos e ferramentas de saúde também utilizam SQLite para gerir registos de saúde com eficiência e segurança [68, 69, 70, 71, 72]. O uso do SQLite não se limita apenas ao domínio móvel, fabricantes de automóveis autónomos também consideram o SQLite para os seus sistemas de gestão de base de dados em veículos, devido à sua confiabilidade e facilidade de integração [73].

Além das aplicações mencionadas, o SQLite também é amplamente adotado em diversas ferramentas de servidor, como em análises de dados em larga escala e em aplicações de bioinformática de alto desempenho. A sua portabilidade e confiabilidade tornam-no uma escolha popular em *datacenters* e ambientes de computação distribuída. Na figura 3 está ilustrada a aplicabilidade da base de dados SQLite [69, 74].

3.1.2 Web Services - Do início ao domínio do REST

Nas últimas três décadas tem-se registado uma proliferação significativa de conjuntos de dados biológicos e de ferramentas para a análise desses dados. Grandes instituições como o **Instituto Europeu de**

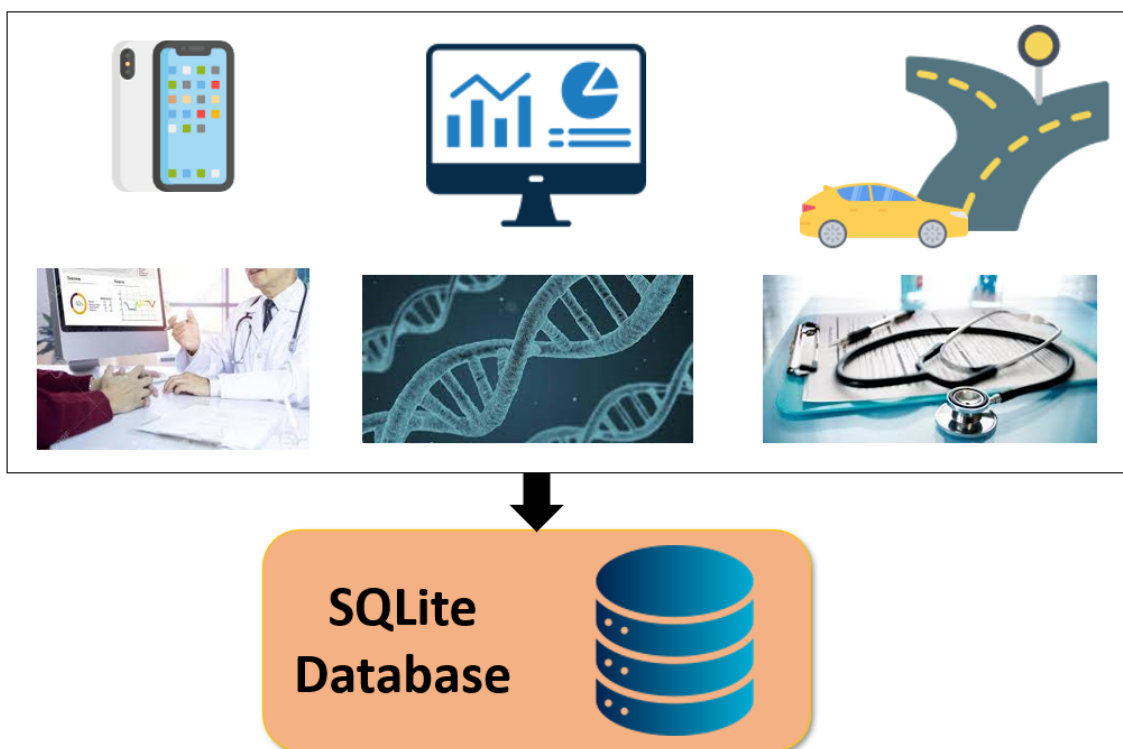


Figura 3: Aplicações da base de dados SQLite. Adaptado de [1].

Bioinformática (EBI) e o **Centro Nacional de Informação Biotecnológica (NCBI)** administram vastos repositórios de dados biológicos e *softwares* especializados [75].

Um problema central atualmente em bioinformática é a acessibilidade a estes conjuntos de dados sem a necessidade de manter cópias locais [76]. As primeiras abordagens exploraram o potencial da Internet, utilizando soluções *peer-to-peer* baseadas em grelhas, que envolviam a colaboração e partilha de recursos entre participantes distribuídos numa rede. Embora se tenham revelado eficazes em redes locais, estas abordagens não se mostraram adequadas em escala global, resultando no que ficou conhecido como ‘problema da interoperabilidade’ [77].

A introdução do protocolo **HTTP** e a popularização da **World Wide Web (WEB)** na década de 1990 representam um avanço significativo, permitindo interfaces **WEB** para acesso a dados e ferramentas de bioinformática. No entanto, esta abordagem não resolveu completamente o problema de interoperabilidade, uma vez que, as interfaces eram autónomas e não comunicavam eficientemente entre si [78].

A solução para este problema veio com o uso de **Interfaces de Programação de Aplicações (APIs)**, servidas sobre **HTTP**, que forneciam respostas facilmente interpretáveis por máquinas. Após breves experiências com diferentes abordagens, tais como o **COBRA**, o padrão **SOAP** foi amplamente adotado. **SOAP** utilizava **XML** para empacotar chamadas de função e dados, mas era criticado pela sua rigidez e complexidade [79].

O **REST**, proposto em 2000 por Roy Fielding [80], emergiu como uma alternativa mais simples e eficiente comparado com a abordagem **SOAP**. Baseado em princípios fundamentais do **HTTP**, o **REST** permite solicitações e respostas **HTTP**, utilizando **URL** para aceder aos recursos. Ao contrário do **SOAP**, não requer bibliotecas especializadas para utilizadores, e as respostas são frequentemente formatadas em **JSON**, proporcionando uma maior flexibilidade e simplicidade. Na década de 2010, **REST** substituiu amplamente o **SOAP** nas **APIs** de bioinformática. A transição para o **REST** trouxe melhorias significativas em termos de flexibilidade, simplicidade e eficiência das interações entre serviços de bioinformática. Até 2020, o **REST** tornou-se o estilo dominante para as **APIs** web de bioinformática, marcando uma evolução significativa na acessibilidade e interoperabilidade desses recursos [81]. Na figura 4 é representada, de forma simplificada, uma arquitetura do tipo **REST**, fornecendo uma visualização clara do funcionamento deste modelo.

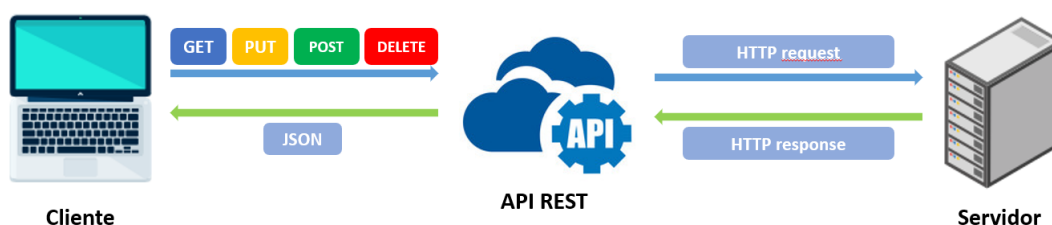


Figura 4: Esquema simplificado de uma arquitetura Rest.

3.1.3 Node.js

O constante avanço da *internet* nos dias de hoje, desafia muitos *sites* a enfrentarem novos problemas, como o aumento das solicitações de múltiplos utilizadores e o crescente número de conexões simultâneas. A linguagem de *script* dinâmica JavaScript tornou-se extremamente popular para utilizadores e é amplamente utilizada no desenvolvimento *web* [82]. O Node.js é uma plataforma baseada no interpretador V8 JavaScript Engine, implementado pelo Google e desenvolvido essencialmente com C e C++, e foi projetado para construir programas de rede orientados a objetos [83]. Utilizando um modelo de E/S sem bloqueio, o Node.js consegue lidar eficientemente com muitas conexões simultâneas, tornando-se especialmente adequado para aplicações em tempo real e com uso intensivo de dados [82].

Uma das características marcantes do Node.js é a sua capacidade de processar solicitações assíncronas de forma eficaz permitindo que um único *thread* atenda a um grande número de clientes. Isto contrasta com abordagens tradicionais de programação que podem bloquear a execução do código enquanto aguardam uma resposta, tornando o Node.js mais leve e eficiente em termos de recursos [82].

O **Node Package Manager (NPM)** desempenha um papel fundamental no ecossistema do Node.js, sendo o maior repositório de pacotes JavaScript disponíveis [84, 85]. Ao permitir a fácil reutilização de código, frequentemente realizada por meio de pacotes de *software*, o **NPM** agiliza significativamente o desenvolvimento de produtos e melhora a qualidade do código. Com acesso a mais de 700 mil pacotes Node.js, disponíveis no repositório online e instaláveis através do comando 'npm', os programadores podem aproveitar uma vasta gama de recursos e funcionalidades prontas para uso [85].

3.1.4 Bibliotecas

React

React é uma biblioteca *front-end* de JavaScript para construção de **Interface de Utilizador (IU)**, inicialmente desenvolvida como uma ferramenta para atender às necessidades internas do Facebook, pelo engenheiro de *software* Jordan Walke [86]. No entanto, não muito depois da sua criação, em 2013, o Facebook tornou o código do React aberto.

Esta biblioteca de renderização foi desenvolvida para proporcionar uma experiência mais dinâmica em *sites* e tornar o processo de desenvolvimento mais eficiente [87]. Este objetivo é alcançado por meio da utilização de componentes de **Interface de Utilizador** reutilizáveis, os quais apresentam dados dinâmicos que podem mudar ao longo do tempo ou em resposta à entrada do utilizador. Os componentes de **IU** reutilizáveis, para além de poderem ser usados novamente em várias partes do código, são também salvos no *cache* da aplicação para que os componentes não precisem de ser renderizados novamente [88]. A seguir estão listados alguns dos motivos que fazem tantos programadores optar por esta biblioteca:

- Facilidade de utilização - por se tratar de uma biblioteca de código aberto, faz com que se trate de uma ferramenta muito simples de ser escrita, compreendida e utilizada;
- Reutilização de componentes - aumentando a produtividade dos programadores;
- Melhor desempenho com Virtual **DOM** - O React torna mais eficiente o processo de atualização do **DOM**, sigla para **Document Object Model**. Nesse sentido, a utilização do Virtual **DOM** reduz os problemas que podem surgir no processo de atualização, visto que é uma etapa que apresenta diversas complicações em projetos direcionados para aplicações *web*. Além disso, a presente ferramenta possibilita a construção de Virtual **DOM** e de outros elementos na memória. Assim, quando ocorre uma mudança no **DOM**, a parte virtual modifica instantaneamente.
- React Native - funciona como uma biblioteca criada para desenvolver aplicações para dispositivos móveis, os quais podem apresentar a configuração de Android ou iOS [89].

3.2 Metodologia da Prova de Conceito

Análise SWOT

A análise **SWOT** passou por várias fases desde a sua concepção. Inicialmente, surgiu na década de 1950 na **Harvard Business School**, onde o professor George Albert Smith Jr usou a técnica para analisar estudos de caso sobre estratégias organizacionais e o seu impacto no ambiente [90, 91, 92]. Algumas fontes apontam para Albert Humphrey, do **Stanford Research Institute**, na década de 1960, como o precursor, utilizando a análise **SWOT** para analisar empresas da *Fortune 500*, que consiste numa lista divulgada, anualmente, com as 500 maiores empresas dos Estados Unidos, com o intuito de desenvolver um novo sistema de gestão e controlo de mudanças [93].

A análise **SWOT** ganhou destaque em 1963 durante uma conferência sobre política empresarial na **HBS**, sendo amplamente discutida e considerada um avanço significativo no pensamento estratégico [94, 95, 96]. Após a década de 1960, a técnica foi adotada por investigadores do planeamento estratégico e foi amplamente reintroduzida na década de 1980 [97, 98]. Por sua vez, nos anos 90, a análise **SWOT** consolidou-se como uma estrutura dominante no campo da gestão estratégica, conforme destacado por [99].

A análise **SWOT**, sigla proveniente da língua inglesa, cuja ilustração está representada na figura 5, é responsável por avaliar as principais forças (*Strengths*), fraquezas (*Weaknesses*), oportunidades (*Opportunities*) e ameaças (*Threats*). Trata-se de uma ferramenta de estratégia empresarial para avaliar a forma como uma organização se compara com a concorrência [100]. Também conhecida como Matriz **SWOT**, foi reconhecida como útil para diferenciar e estabelecer um nicho num mercado mais amplo. Além do contexto empresarial, a análise **SWOT** também pode ser aplicada a nível individual para avaliar a situação de uma pessoa em relação à sua concorrência. A ferramenta inclui considerações internas e externas. As "forças" e as "fraquezas" são fatores internos que refletem vantagens ou desvantagens relativamente à concorrência. Por outro lado, as 'oportunidades' são elementos externos do ambiente que podem ser explorados em benefício da entidade, enquanto as 'ameaças' representam desafios externos que podem dar origem a problemas [101].

A seguir, estão sintetizados os aspetos fundamentais da matriz **SWOT**:

Forças: Dizem respeito às características internas positivas que conferem vantagens competitivas. Desde recursos financeiros robustos até uma equipa altamente qualificada e tecnologia avançada, as for-

	Fatores positivos	Fatores negativos
Fatores internos	S Forças	W Fraquezas
Fatores externos	O Oportunidades	T Ameaças

Figura 5: Ilustração da Matriz representativa das principais características da análise **SWOT**.

ças são os alicerces sobre os quais uma organização pode construir e consolidar uma posição competitiva.

Fraquezas: Representam características internas desfavoráveis que podem representar obstáculos. Identificar problemas na equipa, falta de recursos ou processos operacionais obsoletos são essencial para implementar melhorias significativas.

Oportunidades: Referem-se a fatores externos positivos que uma organização pode explorar. Mudanças de mercado, avanços tecnológicos e tendências sociais representam oportunidades para inovação e crescimento.

Ameaças: São fatores externos adversos que podem desafiar o desempenho da organização. Desde concorrência intensa até mudanças regulatórias, entender e enfrentar ameaças é fundamental para a resiliência organizacional.

Capítulo 4

Caso de Estudo I - Plataforma Colaborativa de Monitorização, em Tempo Real, de análises clínicas, nas Instituições de Saúde

Este capítulo oferece uma visão aprofundada da plataforma *web* desenvolvida, com base na questão de investigação colocada inicialmente. Assim, a plataforma está projetada para aprimorar a monitorização de análises clínicas, nas instituições de saúde. Neste capítulo são destacadas as suas características, funcionalidades e contribuições para o panorama da saúde. Este capítulo explora também como é que esta plataforma, intitulada por '*AnalyticCare*', pode impactar positivamente a eficiência operacional, a tomada de decisões clínicas e, por conseguinte, a qualidade dos cuidados prestados aos pacientes. A '*AnalyticCare*' não é apenas uma ferramenta tecnológica, mas uma solução abrangente para enfrentar desafios específicos no campo das análises clínicas em instituições de saúde.

O código fonte desenvolvido para a elaboração da plataforma *AnalyticCare* pode ser encontrado no GitHub: https://github.com/quiteriap1/AnalyticCare_web_application

4.1 Introdução e contextualização do problema

Num contexto em que a gestão eficiente de dados clínicos é fundamental para aprimorar os cuidados de saúde, surge a necessidade de explorar soluções inovadoras para lidar com conjuntos massivos de informações provenientes de análises clínicas. "Qual o impacto da utilização de uma aplicação de análise de dados no exercício da medicina e na melhoria da qualidade de vida dos utentes?" foi a questão colocada inicialmente e para melhor responder a esta pergunta, foi desenvolvida esta aplicação. Assim, este capítulo concentra-se na implementação e análise de uma plataforma diligente, ao qual foi dado o nome de *AnalyticCare* e que será explicada melhor à posteriori. Esta plataforma recebe e processa um extenso *dataset* contendo perto de dois milhões de dados de análises clínicas gerados de forma sintética

para fornecer uma representação simulada e controlada de diferentes cenários existentes em instituições de saúde. A importância desta plataforma revela-se na sua capacidade de transformar dados brutos em *insights* clinicamente relevantes. Particularmente, a *AnalyticCare* dispõe visualizações gráficas que facultam aos médicos uma análise abrangente e detalhada da variação nos valores do paciente para cada exame, ao longo do tempo. Este recurso é essencial para a compreensão da evolução clínica individual, permitindo ajustes precisos nos planos de tratamento. Contudo, o desafio não reside apenas na gestão deste conjunto vasto de dados, mas na habilidade de oferecer aos profissionais de saúde ferramentas que vão além da simples visualização temporal. É aqui que a *AnalyticCare* se destaca, pois proporciona ao médico a capacidade de comparar os valores de um paciente específico com os de outros pacientes, que se encontram na mesma faixa etária ou provenientes da mesma localidade. Esta funcionalidade procura enriquecer a análise clínica, ao considerar fatores demográficos e geográficos que podem influenciar o estudo dos resultados dos exames, por exemplo, ao auxiliar os médicos a identificarem padrões, tendências e anomalias de forma eficiente. Adicionalmente, a plataforma exibe alertas, sempre que se verifiquem análises clínicas com valores fora do limite considerado adequado para a designação em específico. Ou seja, sempre que um médico procure por um paciente em específico e este tenha no seu histórico de análises um valor fora do considerado normal, na página principal do paciente surgirá um alerta sobre esse valor. Esta funcionalidade proporciona uma vigilância proativa e facilita intervenções precoces quando necessário.

4.2 Análise de Requisitos

A análise de requisitos desempenha um papel central na conceção e implementação bem-sucedida da plataforma *AnalyticCare*. Nesta secção, serão detalhados os requisitos-chave identificados durante o processo de análise, os quais guiaram o desenvolvimento desta aplicação *web* para a gestão e interpretação de dados clínicos.

1. Gestão eficiente de grandes *datasets*

Um requisito fundamental foi a capacidade da plataforma de lidar eficientemente com um grande volume de dados, dado o extenso *dataset* composto por mais de dois milhões de registos.

2. Visualização temporal detalhada

Para atender às necessidades dos profissionais de saúde, a plataforma foi projetada para oferecer visualizações temporais detalhadas, permitindo aos médicos examinar de forma minuciosa

a variação dos valores de cada paciente para cada exame, ao longo do tempo. Isto implicou o desenvolvimento de gráficos interativos e ferramentas de navegação intuitivas.

3. Comparação entre pacientes e grupos demográficos

Uma funcionalidade crucial foi a capacidade de comparar os valores de um paciente específico com os de outros na mesma faixa etária ou provenientes da mesma localidade. Isto exigiu a incorporação de filtros que permitissem uma análise contextualizada.

4. Alerta de valores críticos

Além das funcionalidades até agora mencionadas, a plataforma *AnalyticCare* foi também projetada para emitir alertas aos profissionais de saúde sempre que sejam identificados valores fora dos limites recomendados, para cada análise clínica em específico.

5. Segurança e privacidade

Considerações rigorosas de segurança e privacidade foram incorporadas, assegurando que os dados sensíveis dos pacientes fossem protegidos em conformidade com os regulamentos de saúde e padrões de segurança.

Ao atender a estes requisitos, a *AnalyticCare* foi moldada para ser uma ferramenta robusta, intuitiva e altamente funcional, oferecendo aos profissionais de saúde uma abordagem avançada na análise de dados clínicos e na tomada de decisões informadas.

4.3 Origem do dataset

Os dados utilizados por esta aplicação foram gerados de forma sintética para fornecer uma representação simulada e controlada de diferentes cenários. Em vez de depender de conjuntos reais, optou-se por criar artificialmente informações que abrangem uma variedade de casos, permitindo testar e otimizar o desempenho da aplicação em condições diversificadas. Esta abordagem proporciona flexibilidade e controle sobre os dados de entrada, assegurando que a aplicação possa lidar eficazmente com situações diversas, enquanto preserva a privacidade e a confidencialidade de informações sensíveis. Posteriormente, os dados foram submetidos a um processo de pré-processamento, utilizando a linguagem de programação Python no ambiente Jupyter Lab. A biblioteca Pandas foi utilizada e, durante esta etapa, os dados foram carregados para um *dataframe*, que se trata de uma estrutura de dados eficiente para manipulação e

análise de conjuntos de dados tabulares. O uso da biblioteca Pandas permitiu a aplicação de transformações necessárias, como limpeza e seleção de dados, garantindo que estes estivessem prontos para serem consumidos pela aplicação *web* desenvolvida. Esta abordagem proporcionou flexibilidade e eficiência no tratamento dos dados, contribuindo para a robustez e desempenho da aplicação.

A seguir, segue uma breve explicação de todas as colunas constituintes do *dataset* utilizado para a elaboração da aplicação *web*:

EXAME: Nesta coluna, são registados números que servem como códigos distintos associados a uma variedade de exames clínicos conduzidos no hospital. Cada código identifica um teste específico, abrangendo uma gama diversificada de procedimentos analíticos, tais como 'Hemograma', 'Glicemia', entre outros. Normalmente, um exame específico é realizado como parte de um episódio clínico mais amplo. Estes códigos proporcionam uma forma eficiente de categorizar e diferenciar os tipos particulares de exames realizados, facilitando a organização e interpretação dos dados clínicos.

DESIGNAÇÃO: Esta coluna fornece detalhes adicionais sobre o exame específico referenciado na coluna 'EXAME'. Cada entrada nesta coluna identifica uma análise específica ou um componente de exame, detalhando o que está a ser avaliado num procedimento mais constrito. Por exemplo, para o exame 'Hemograma' podem ser analisadas diferentes designações, tais como: 'neutrófilos', 'basófilos', 'leucócitos'. Ou seja, para cada paciente, um mesmo exame contém diferentes designações.

VALOR: Refere-se aos resultados dos exames clínicos, podendo ser expressos com valores numéricos, ou qualificados como 'positivo' ou 'negativo'. Nesta coluna, são registados os dados resultantes dos procedimentos analíticos, abrangendo tanto medições quantitativas quanto indicações qualitativas.

UNIDADE: Indica a unidade de medida associada aos valores registados na coluna 'VALOR'. Por exemplo, miligramas por decilitro (mg/dL) para a glicose ou milimoles por litro (mmol/L) para valores de potássio.

LIMITE: Representa os limites de referência associados a cada exame. Estes limites indicam os intervalos considerados normais para os resultados dos exames, permitindo a avaliação dos valores obtidos.

EPISÓDIO: Nesta coluna encontram-se registados números que estão associados a diferentes episódios clínicos na trajetória de cuidados de um paciente. Cada número pode ser usado para agrupar

exames relacionados a eventos ou períodos específicos de cuidado. Estes números referentes aos episódios desempenham um papel fundamental na organização e compreensão do histórico clínico do paciente, permitindo a categorização sistemática de exames associados a diferentes fases do tratamento ou a eventos particulares ao longo do tempo. Desta forma a coluna 'EPISÓDIO' facilita a análise contextualizada e contínua dos dados clínicos, fornecendo dados valiosos sobre a evolução da condição de um paciente em resposta ao tratamento e aos eventos clínicos.

DATA_NASCIMENTO: Como o próprio nome diz, refere-se à data de nascimento do paciente submetido aos exames clínicos. Esta informação é relevante para a contextualização dos resultados em relação à idade do paciente. Com esta informação, o médico pode comparar os valores de um certo paciente com os valores de outros pacientes da mesma idade e tirar conclusões.

DATA_EXAME: Diz respeito ao dia em que foi realizado o exame. Esta informação é crucial para o acompanhamento do histórico clínico do paciente e para a análise temporal dos resultados.

SEXO: Esta coluna serve para termos informações sobre o sexo do paciente submetido aos exames clínicos. Esta informação pode ser útil para considerações específicas de gênero em relação aos resultados dos exames.

LOCALIDADE: Fornece informações sobre a localidade do paciente. É muito importante pois permite ao profissional de saúde comparar os valores de um certo paciente com os valores de outros pacientes da mesma localidade.

NUM_SEQUENCIAL: Corresponde ao número sequencial único atribuído a cada paciente. Este identificador exclusivo é único para cada paciente e tem a função essencial de ser um marcador permanente que acompanha o paciente ao longo da sua jornada de cuidados no hospital. Desta forma, o número sequencial é utilizado para rastrear e organizar sistematicamente todos os exames realizados por um paciente específico.

4.4 Arquitetura

Desenvolver uma aplicação é um desafio que demanda tempo e envolve um processo intrincado. É necessário um investimento considerável de tempo até que a aplicação evolua para um nível em que apresente uma interface coesa e um sistema atraente e intuitivo para os utilizadores. A arquitetura da plataforma precisa de ser construída com uma estrutura modular, garantindo escalabilidade e facilitando ajustes posteriores de forma ágil.

O presente caso de estudo centra-se na elaboração de uma plataforma colaborativa de monitorização de análises clínicas, em instituições de saúde, com o objetivo de avaliar o impacto que a utilização de aplicações de análises de dados no âmbito da medicina podem ter, relativamente à melhoria da qualidade de vida dos pacientes. Posto isto, foi então desenvolvido o artefacto, cuja arquitetura se ilustra na figura 6, com o propósito de aprimorar o suporte às decisões dos profissionais de saúde. A plataforma consome dados de uma base de dados que está hospedada num servidor local. Para solicitar essas informações foi criado um *web Service* em Node.js. O seu funcionamento é bastante simples e inicia-se com o pedido por parte do cliente. Após a conexão à base de dados ser bem-sucedida, é possível realizar as queries, em SQL para que sejam selecionados os dados desejados. Por último, o *web Service* transforma os dados obtidos no formato JSON e estes são enviados para o cliente. É importante referir que todos os gráficos foram efetuados em Recharts que facilita a construção dos gráficos por ser uma combinação do React com o D3 (biblioteca de visualização de JavaScript).

Foram construídos alguns endpoints criteriosamente projetados para atender às necessidades específicas da aplicação desenvolvida. Entre eles, destacam-se dois do tipo 'Post', proporcionando a capacidade de submeter dados para a aplicação, enquanto os restantes oito endpoints são do tipo 'Get', ou seja, possibilitam a consulta eficiente de informações. A seguir estão apresentados os endpoints usados:

Endpoints do tipo 'Post': '/signup', '/login'

Endpoints do tipo 'Get': '/utentes', '/exames/:num_sequencial', '/utentes-por-exame', '/exames-por-utente-designacao', '/informacoes-gerais-exames-localidades', '/informacoes-gerais-exames-idades', '/utentes-por-localidade-designacao', '/utentes-por-idade-designacao'

4.4.1 Nome da aplicação: **AnalyticCare**

O nome da plataforma resulta da união de duas palavras diretamente relacionadas com a mesma. O termo '*Analytic*' reflete um compromisso com a precisão na análise de dados clínicos. A plataforma foi desenvolvida para fornecer análises detalhadas e confiáveis para melhorar a tomada de decisões médicas. O termo '*Care*', que em português significa 'Cuidado' foi pensado por estar, diretamente, associado ao setor da saúde e, por outro lado, por todo o cuidado que foi colocado no desenvolvimento da plataforma, onde cada detalhe, foi pensado ao pormenor para que dela resultassem não apenas funcionalidades analíticas avançadas, mas uma ferramenta integral e confiável que se alinha perfeitamente com as necessidades e exigências do ambiente clínico. Cada aspeto da *AnalyticCare* foi meticulosamente elaborado para garantir não apenas precisão nos dados, mas também uma experiência de usuário fluida e centrada no cuidado,

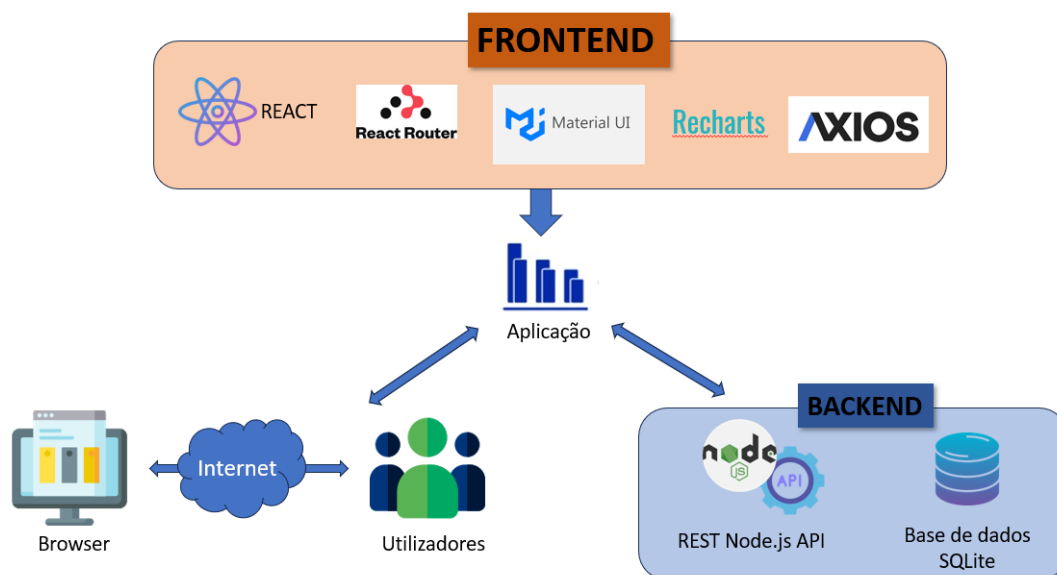


Figura 6: Arquitetura da aplicação *web* desenvolvida.

refletindo assim a dedicação à excelência na monitorização de análises clínicas e na promoção contínua da qualidade nos serviços de saúde. O nome da plataforma ao unir estes dois elementos, simboliza a fusão harmoniosa entre a análise de dados precisa e o compromisso com o cuidado, no âmbito da saúde digital. Na figura 7, está representado o logótipo elaborado para a aplicação e que de forma se destaca na mesma.



Figura 7: Logótipo da aplicação *web* desenvolvida.

4.4.2 Página Sign In e página Logout

As páginas de *Sign In* e *Logout* desempenham papéis essenciais na segurança, acessibilidade e controlo de utilizadores numa aplicação. São importantes para garantir que apenas pessoas com autorização tenham acesso à aplicação.

Create Account e Sign In

A primeira página com a qual se depara o médico assim que acede à plataforma, pela primeira vez, é a página de '*Create Account*'. Nesta página são-lhe pedidas algumas informações pessoais de forma a ficar registado na plataforma. A partir daqui, sempre que o médico que já efetuou o registo na aplicação se voltar a tentar conectar, só precisa de passar pela página de '*Login*', e inserir o seu email e password,

definidos previamente. Na figura 8, estão representadas as duas páginas de acesso à aplicação.

The image displays two web forms side-by-side. The left form, titled 'Login', features a purple lock icon above the title, followed by input fields for 'Email Address *' and 'Password *'. Below these fields is a blue button labeled 'LOGIN' and a link that reads 'Don't have an account? Sign Up'. The right form, titled 'Create Account', also has a purple lock icon above the title. It includes input fields for 'First Name *', 'Last Name *', 'Email Address *', and 'Password *'. Below these fields is a blue button labeled 'CREATE ACCOUNT' and a link that reads 'Already have an account? Sign in'.

Figura 8: Páginas iniciais com sistema de segurança, onde o médico efetua o seu registo na aplicação (esquerda) e, onde insere os seus dados sempre que acede à aplicação (direita).

Logout

Por sua vez, o botão 'logout' oferece aos utilizadores a opção de encerrar a sessão na plataforma. Esta funcionalidade é vital para garantir a segurança e a privacidade do profissional de saúde, assim como dos dados dos pacientes.

4.4.3 Página inicial

Após concluírem o registo na aplicação, os médicos são direcionados para a página principal da aplicação, que está representada na figura 9. Trata-se do ponto central e é onde os utilizadores têm acesso às principais funcionalidades da plataforma. A seguir, está uma explicação detalhada de todos os elementos presentes.

Cabeçalho

O cabeçalho, localizado no canto superior da página, exibe o logótipo e o nome da aplicação, o que proporciona uma identidade visual única. Ao lado, encontram-se dois botões que servem como atalhos para aceder a duas páginas específicas. O primeiro botão direciona para a página de tendências por idade, enquanto o segundo orienta para a página de tendências por localidade. Estes atalhos oferecem uma navegação rápida e eficiente para áreas específicas da aplicação e serão abordados, detalhadamente, no decorrer deste capítulo. Além disso, no canto direito do cabeçalho, encontra-se o botão de

Num Sequencial	Data Nascimento	Sexo	Localidade	Ação
100012	1970-08-01	Feminino	TINALHAS	→
100027	1960-10-07	Feminino	CASTELO-BRANCO	→
100034	1991-11-02	Feminino	CASTELO-BRANCO	→
100035	1969-08-07	Feminino	CASTELO-BRANCO	→
100036	1930-11-11	Feminino	CASTELO-BRANCO	→
100051	1965-04-17	Feminino	CASTELO-BRANCO	→
100066	1986-10-02	Feminino	PEDRA-DO-ALTAR	→
100068	1987-08-18	Feminino	CIDREIRO	→
100084	1989-11-15	Feminino	SÃO-MIGUEL-D' ACHA	→
100091	1983-09-11	Feminino	CASTELO-BRANCO	→
100106	1939-04-18	Feminino	PROENÇA-A-NOVA	→
100111	1930-05-30	Feminino	LOURICAL-DO-CAMPO	→
100116	1935-11-09	Feminino	ORVALHO	→
100122	1954-10-30	Feminino	ALCAINS	→
100125	1923-05-14	Feminino	TABERNA-SECA	→
100130	1965-11-17	Feminino	ROCHAS-DE-CIMA	→
100136	1931-05-26	Feminino	CASAL-SANTANA	→

Figura 9: Página principal da plataforma que surge após registo na aplicação.

logout, permitindo ao utilizador encerrar a sua sessão de forma controlada e garantindo a segurança da sua conta.

Tabela principal

A tabela principal é o componente central da página, apresentando dados essenciais sobre os pacientes. É possível ordenar a tabela por qualquer coluna, como número sequencial, data de nascimento, sexo e localidade, proporcionando aos utilizadores a flexibilidade de organizar os dados conforme as suas necessidades específicas. Uma funcionalidade crucial da tabela é a inclusão da coluna 'Ação', onde cada linha desta coluna é composta por uma 'seta'. Ao clicar nesta 'seta', o médico é direcionado para outra página que contém informações detalhadas sobre um paciente em específico. É importante referir que, como já foi explicado à priori, cada paciente é identificado por um número sequencial exclusivo, simplificando a identificação e referência.

Dada a quantidade significativa de dados, a tabela é dividida em várias páginas, inicialmente exibindo 25 dados por página. No entanto, para acomodar as preferências e necessidades do médico, a opção de modificar a quantidade de linhas exibidas por página é oferecida, permitindo escolher entre 15, 25, 50 e 75 linhas por página.

4.4.4 Página de detalhe do paciente

A página de detalhe do paciente, destaca-se como uma ferramenta de extrema importância para os médicos, fornecendo *insights* detalhados e personalizados que desempenham um papel crucial na tomada de decisões clínicas. A sua divisão em três partes distintas, como podemos ver na figura 10, reflete a abordagem metódica para oferecer uma visão completa e aprofundada do histórico de saúde do paciente.

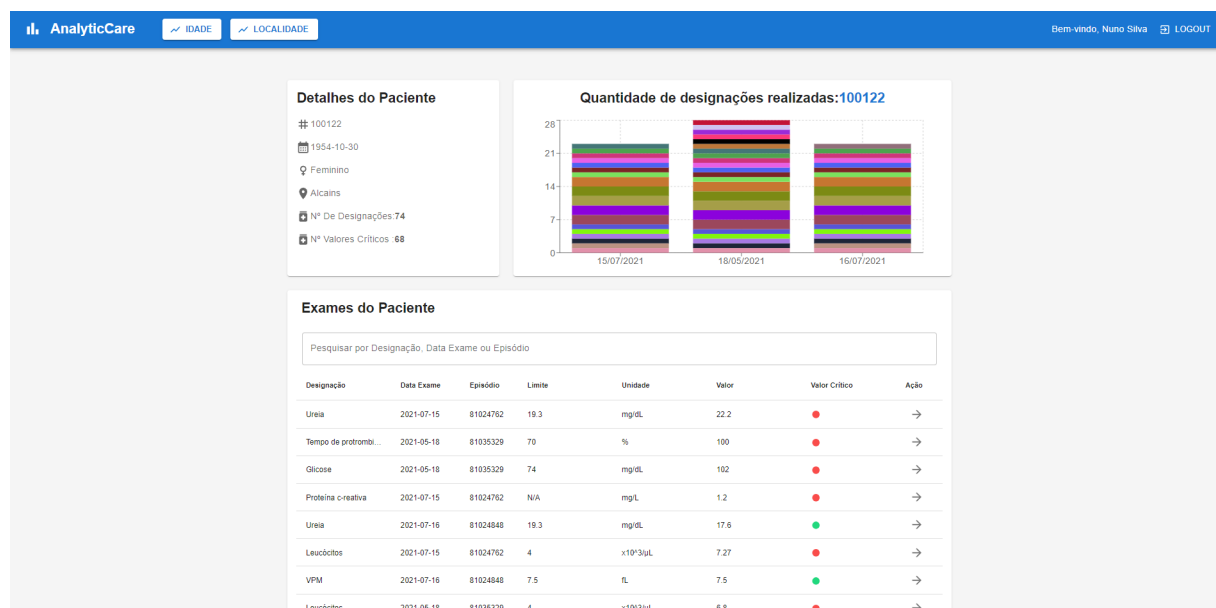


Figura 10: Página de detalhe do paciente.

Detalhes do paciente

No canto superior esquerdo da página, encontra-se um espaço dedicado a reunir todas as informações gerais do paciente, incluindo detalhes essenciais como o número sequencial, data de nascimento, sexo, localidade do paciente e histórico de análises clínicas, tais como o número de designações que foram realizadas pelo paciente, ao longo do tempo e também valores críticos para o paciente. Esta última parte é uma das mais cruciais da plataforma, visto servir como um alerta assim que o médico abre a página de detalhe do paciente. Refere-se ao número de designações cujos valores estão fora dos limites recomendados, podendo demonstrar algum risco para o paciente e, consequentemente, indicar que precisa de algum tipo de vigilância. Esta visão abrangente é fundamental para compreender o contexto médico do paciente de forma rápida e eficaz e pode analisar-se com mais pormenor, na figura 11.

Número total de designações realizadas pelo paciente

Através de um gráfico de barras, com o aspeto do presente na figura 12, o médico tem um resumo do número total de exames realizados pelo paciente que está a estudar e, ao passar o rato por cima

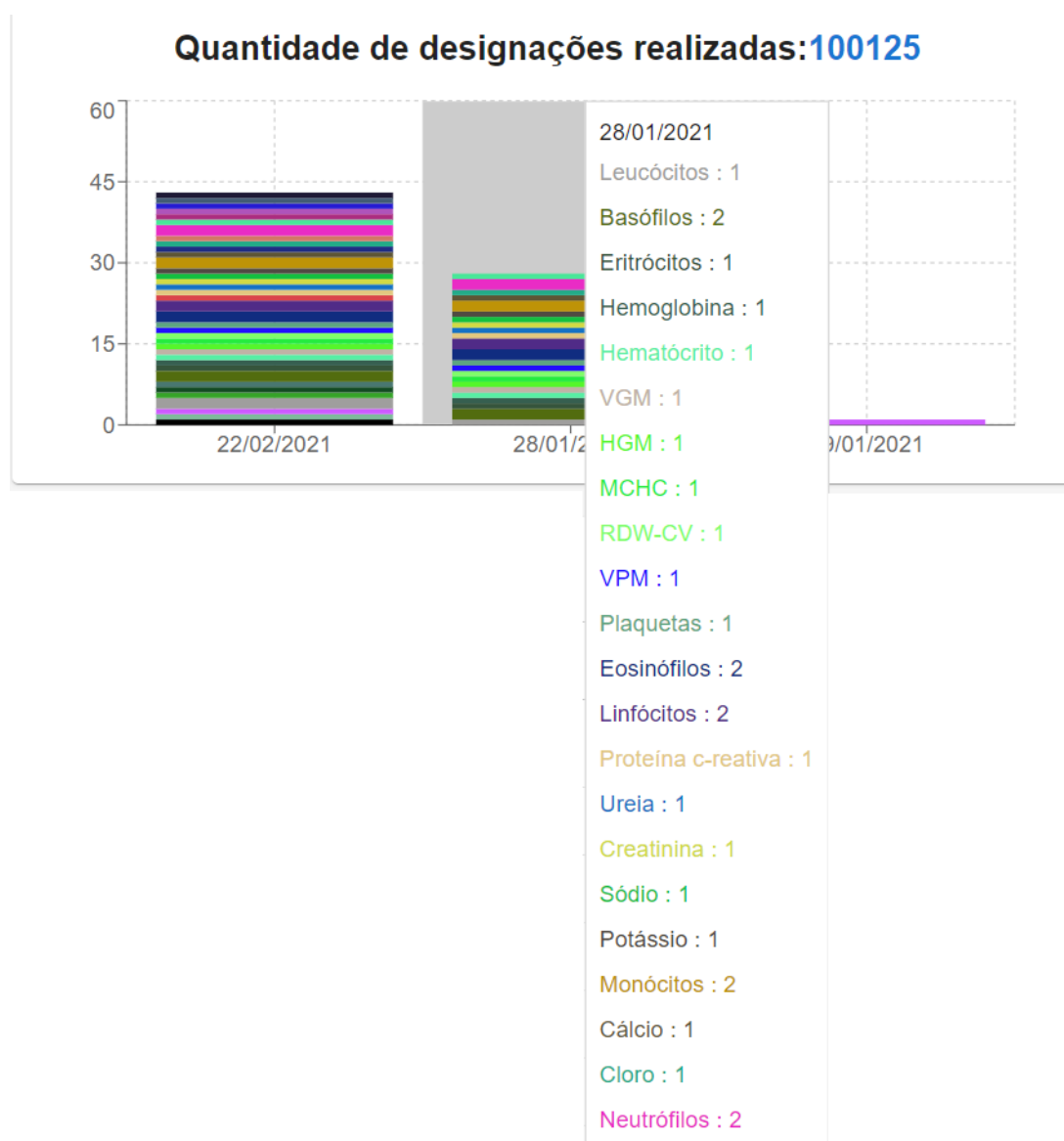


Figura 12: Quantidade de designações analisadas, ao longo dos exames realizados pelo paciente.

Evolução dos valores de cada designação, ao longo do tempo

Seguindo o raciocínio anterior, ao clicar no botão de 'Ação' o médico é direcionado para uma página exclusiva que apresenta um novo gráfico de barras. Na figura 13, está representado um gráfico referente aos valores dos 'leucócitos', ao longo de três exames, para um certo paciente, e é um exemplo do que o utilizador da *AnalyticCare* pode observar. Este gráfico oferece uma perspetiva detalhada das variações nos valores de uma mesma designação ao longo do tempo, concentrando-se numa única análise. Este gráfico combina todos os valores associados à designação em questão, obtidos em diferentes exames realizados pelo paciente. Desta forma, o médico pode examinar de maneira abrangente como os resultados desta designação evoluíram ao longo de diferentes exames, proporcionando *insights* valiosos para a análise clínica. Esta funcionalidade facilita o acompanhamento contínuo do paciente ao longo do tempo,

permitindo ao médico analisar e comparar variações nos resultados das análises. Esta abordagem auxilia também na tomada de decisões informadas e permite ao profissional de saúde um tratamento mais adequado e preciso ao paciente. Para além disso, no gráfico está também indicado o limite esperado para os resultados da designação em específico. A seguir são apresentados alguns exemplos de cenários possíveis, com o objetivo de ilustrar a importância do gráfico no estudo das análises clínicas:

- **Monitorização eficaz do tratamento:** Um paciente está em tratamento para uma condição específica e realiza análises regulares. O gráfico permite ao médico observar se os valores da designação relacionada à condição estão a progredir conforme o esperado. Se os resultados indicarem uma melhoria ao longo do tempo, isso sugere que o tratamento está a ser eficaz. No entanto, caso os resultados mostrem uma estagnação ou piora, isso pode indicar a necessidade de ajustar a terapia atual, considerar intervenções adicionais ou reavaliar o plano de tratamento em conjunto com o paciente.
- **Contextualização de valores fora do normal:** Um paciente apresenta valores fora dos limites considerados normais para uma determinada análise. O gráfico possibilita ao médico comparar esses valores com outras designações ao longo dos exames. Se outros resultados estiverem consistentes e explicáveis, isso pode indicar que esses valores atípicos são normais para aquele paciente em específico, fornecendo uma visão mais abrangente da sua saúde.
- **Monitorização a longo prazo:** Ao estudar o gráfico, o médico pode identificar tendências a longo prazo nas análises de um paciente. Isto é especialmente útil para condições crónicas. Se houver flutuações consistentes em certas designações ao longo dos exames, o médico pode ajustar o plano de tratamento conforme necessário.

Ao recorrer a estes gráficos para verificar de que forma variam os valores das designações analisadas, o médico pode concluir o que é mais indicado para cada paciente. O profissional de saúde pode deparar-se com vários cenários e é sempre importante ter em consideração os valores de outras designações, para conseguir retirar as melhores conclusões.

4.4.6 Página de tendência por idade

Esta página exibe um gráfico com a tendência por idade e é de extrema importância para os médicos, proporcionando uma ferramenta valiosa para a análise aprofundada das análises clínicas. Nesta página, representada na figura 14, o médico pode selecionar a faixa etária desejada, com intervalos de 5 anos e,

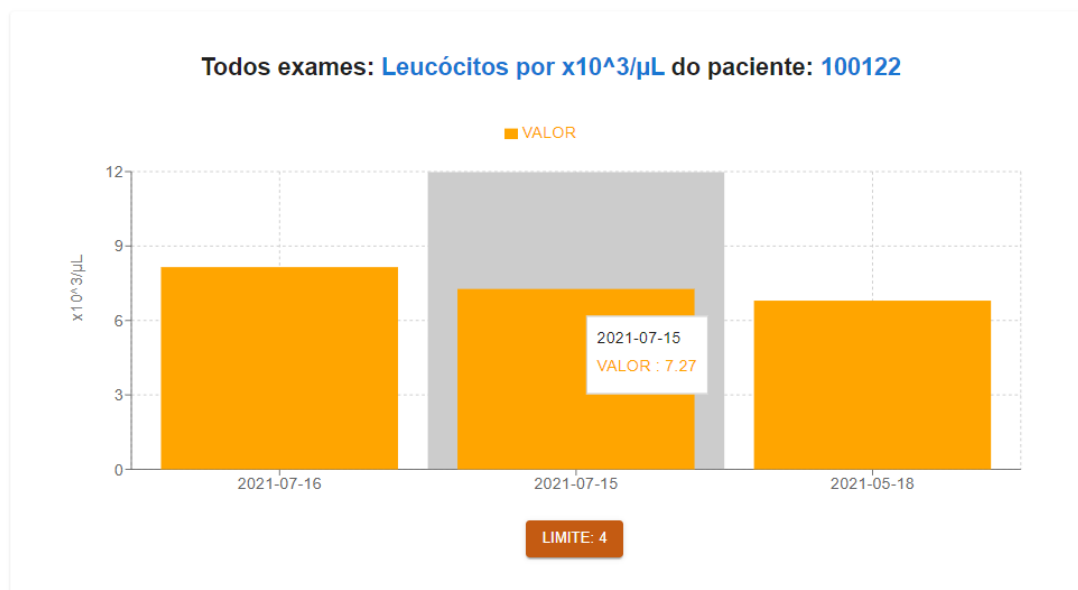


Figura 13: Variação de todos os valores de certa designação, realizados em três exames distintos.

escolher uma designação específica para a análise, com base no que está a estudar. Ao permitir que o médico escolha a faixa etária juntamente com a designação de interesse, a plataforma possibilita a identificação de padrões, tendências e variações nos resultados. Desta forma, o profissional de saúde pode tirar conclusões acerca do paciente que está a analisar em detalhe e perceber se os valores dele para determinada designação são muito díspares quando comparados com os valores de outros pacientes, da mesma faixa etária, ou se seguem a tendência global. O médico pode também servir-se desta ferramenta para personalizar tratamentos, antecipar tendências associadas a faixas etárias específicas, por exemplo, ajuda na identificação precoce de padrões, e ajustar estratégias de cuidados de saúde conforme necessário.

Na figura 14 verificam-se alguns valores significativamente mais elevados em comparação com o restante dos valores. Estes valores destacados podem ser indicativos de pacientes com resultados anormais, possivelmente relacionados a condições de saúde específicas que estão a ser tratadas. A capacidade do médico de selecionar uma faixa etária específica e uma designação para análise permite um estudo mais preciso destes casos. É importante ressaltar que, embora estes valores se destaquem, é crucial considerar cuidadosamente o contexto clínico dos pacientes, pois podem representar tanto condições específicas quanto características normais para esses pacientes.

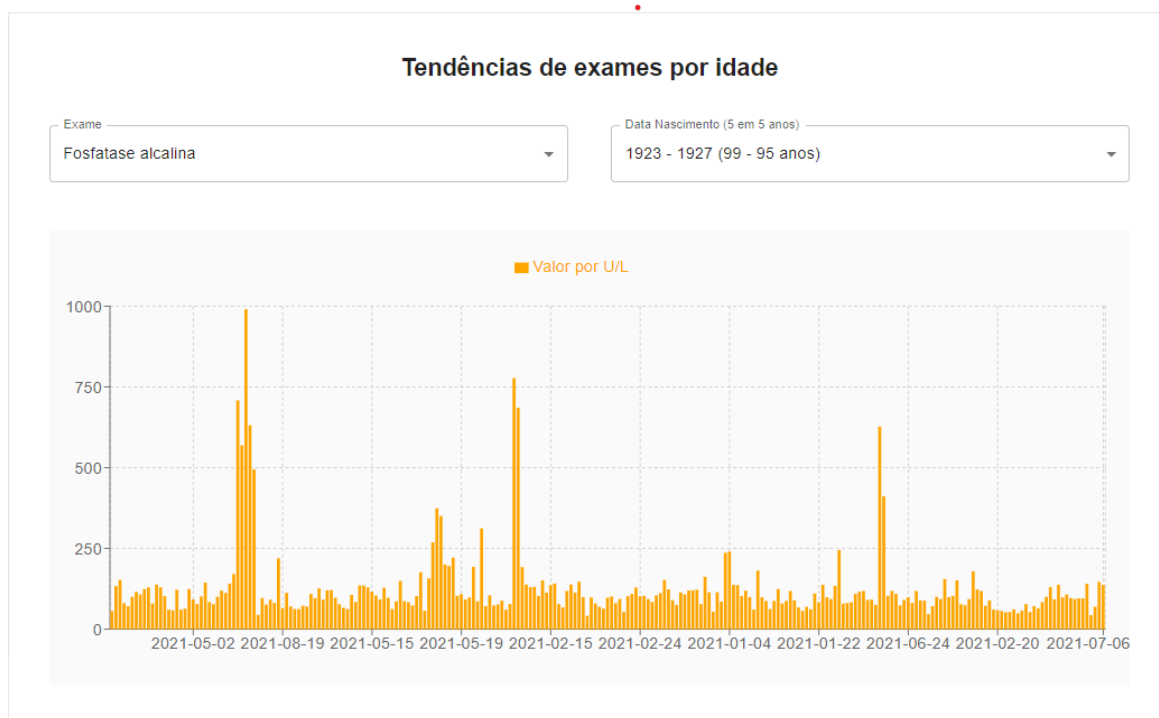


Figura 14: Tendência por idade.

4.4.7 Página de tendência por localidade

À semelhança da página anterior, esta refere-se à tendência por localidade e está representada na figura 15. É igualmente crucial para uma análise bem fundamentada para os profissionais de saúde, ajudando a tirar conclusões com base na localização geográfica dos pacientes. Nesta página, o médico pode escolher uma localidade específica e uma designação, com base no que está a analisar. Esta funcionalidade, ao gerar um gráfico onde estão reunidos todos os valores para a designação selecionada, provenientes de pessoas da mesma localidade, permite ao utilizador entender variações geográficas nos resultados e identificar possíveis correlações entre a localização dos pacientes e os valores das análises. Esta tendência revela-se muito importante porque pode identificar problemas em comum com vários pacientes da mesma localidade, por exemplo, algum problema devido à ingestão de algum alimento contaminado, produzido e consumido na zona em específico. Assim, com este gráfico é fácil decifrar quais as análises que sofreram alterações e permite aos profissionais de saúde agir rápido com toda a população.

Além disso, o gráfico de tendência por localidade pode ser especialmente valioso ao investigar surtos de doenças específicas numa determinada região. Por exemplo, se houver um aumento repentino nos valores de certas análises clínicas em pacientes de uma localidade específica, isso pode indicar a presença de um surto de doença contagiosa. A capacidade de visualizar rapidamente essas variações geográficas permite aos profissionais de saúde tomar medidas preventivas, implementar estratégias de saúde pública

e direcionar recursos para áreas afetadas. Desta forma, o gráfico não apenas contribui para uma análise aprofundada das condições de saúde locais, mas também serve como uma ferramenta proativa na gestão de surtos e na proteção da saúde da comunidade.



Figura 15: Tendência por localidade.

4.5 Resultados e discussão

A elaboração da plataforma de monitorização de análises clínicas revela uma série de resultados e implicações importantes para o contexto da assistência médica. A eficiência na monitorização é notável, proporcionando aos profissionais de saúde uma ferramenta acessível e eficaz para visualizar, interpretar e acompanhar os resultados clínicos. A navegação intuitiva e eficiente da plataforma desempenha um papel crucial, garantindo que os médicos possam aceder facilmente a informações pertinentes, otimizando assim o tempo dedicado ao estudo das análises clínicas.

A introdução de funcionalidades de análise por faixa etária e localidade evidencia uma abordagem personalizada na interpretação dos resultados. Isto permite aos médicos filtrar os dados de maneira mais específica, o que contribui, significativamente, para a compreensão mais profunda dos padrões e tendências, possibilitando uma análise clínica mais refinada. Além disso, a capacidade de realizar comparações diretas com as tendências de faixa etária e localidade distinguem a plataforma. Isto significa que os médicos não estão restritos à análise individual de cada paciente, pois agora têm a capacidade de con-

textualizar e comparar os resultados com dados demográficos mais amplos, oferecendo uma visão mais holística e eficiente do panorama clínico. Outra vantagem que a aplicação apresenta é também o facto de a plataforma avisar o profissional de saúde quando os valores do paciente não estão de acordo com o esperado, ou seja, quando ultrapassam ou ficam aquém dos valores esperados para cada designação. Ao avisar o médico de possíveis anomalias nos valores das análises, podem ser implementadas intervenções precoces e personalizadas. Esta funcionalidade não apenas agiliza a identificação de situações críticas, mas também aprimora a capacidade de resposta do profissional de saúde, permitindo a tomada de medidas preventivas de forma proativa. Esta abordagem preventiva pode resultar em melhorias significativas na qualidade do atendimento e na eficácia dos protocolos de tratamento.

Capítulo 5

Prova de conceito

Após a conclusão do desenvolvimento da plataforma, é fundamental validar a sua aplicabilidade num ambiente e contexto reais. Com este propósito, foi conduzida uma **Prova de Conceito (PoC)** com o objetivo de demonstrar a adaptabilidade da tecnologia desenvolvida às necessidades específicas da área de Tecnologia de Informação.

Neste contexto, foi conduzida uma avaliação do protótipo desenvolvido, com o objetivo de validar a utilidade da ferramenta de apoio à decisão e à prática clínica desenvolvida, através de uma análise ***Strenghts, Weaknesses, Opportunities and Threats (SWOT)***.

Análise **SWOT**

A teoria aplicada a seguir foi meticulosamente elaborada com base nas definições da análise ***Strenghts, Weaknesses, Opportunities and Threats (SWOT)***, proporcionando uma compreensão abrangente das forças, fraquezas, oportunidades e ameaças relativas ao artefacto desenvolvido. Detalhes aprofundados podem ser consultados na secção 3.2 do capítulo 3, onde a análise **SWOT** é minuciosamente abordada, oferecendo *insights* fundamentais para a compreensão do quadro metodológico.

Forças:

- Apresentação da informação de forma resumida e objetiva na ótica dos profissionais de saúde.
- Segurança dos dados através de um mecanismo de autenticação da plataforma.
- A capacidade de análise por faixa etária e localidade destaca a aplicação, permitindo uma abordagem personalizada na interpretação de dados.
- A funcionalidade de alertar profissionais de saúde sobre valores fora do esperado contribui para intervenções precoces e personalizadas.
- A inclusão de gráficos de barras oferece uma representação visual eficaz, facilitando a compreensão

dos resultados.

- A aplicação demonstra eficácia na monitorização de análises clínicas, fornecendo uma ferramenta acessível e eficiente para profissionais de saúde, através de um sistema evolutivo e modular, podendo ser melhorado continuamente, com base na estrutura criada inicialmente.

Fraquezas:

- Visto que a aplicação exige uma conexão constante com a internet, podem surgir desafios em áreas com conectividade limitada.
- Quantos mais dados a aplicação consumir, mais lenta esta se pode tornar.

Oportunidades:

- Aumentar a qualidade e eficiência da plataforma com a utilização de novas tecnologias.
- Aumento da qualidade e eficácia da tomada de decisões médicas, relativamente a análises clínicas, recorrendo a tecnologias inovadoras.
- Oportunidade de expandir a aplicação com a inclusão de novas tendências, para além das já existentes, tendência por idade e por localidade, permitindo ao médico tomar decisões mais globais.

Ameaças:

- Dificuldade em adaptar a plataforma ao dia a dia, por parte dos profissionais de saúde, visto que a plataforma exigirá uma mudança e adaptabilidade por parte de cada profissional.
- Aparição de novas ferramentas, ocorrendo assim concorrência e competitividade com a interface desenvolvida.

Capítulo 6

Conclusões e Trabalho Futuro

O projeto descrito ao longo da dissertação teve como principal foco estudar de que forma a utilização de uma plataforma de análise de dados clínicos se revela uma mais valia no ramo da medicina e de que modo pode melhorar a qualidade de vida dos pacientes. O estudo de análises clínicas envolve uma observação precisa e meticulosa, pois os resultados desempenham um papel crucial nas decisões médicas, influenciando diagnósticos, tratamentos e prognósticos. A confiabilidade e a precisão dessas análises são fundamentais para assegurar a qualidade dos cuidados de saúde prestados aos pacientes.

Diante da complexidade e diversidade dos dados gerados pelas análises clínicas, evidenciou-se a necessidade de uma abordagem inovadora na sua monitorização. A incorporação de ferramentas avançadas para análise e interpretação destes dados revelou-se crucial, assegurando a sua conformidade com os mais elevados padrões de qualidade. A importância destas análises vai além da mera obtenção de resultados. Estas desempenham um papel crucial na compreensão das condições de saúde dos pacientes, orientando uma abordagem personalizada e eficaz.

Num cenário onde a tecnologia desempenha um papel cada vez mais crucial na assistência médica, a plataforma desenvolvida vem dar resposta à questão de investigação colocada no início do desenvolvimento do presente projeto, destacando-se como uma contribuição significativa para a monitorização eficaz das análises clínicas. Ao centrar-se no paciente e na otimização das práticas clínicas, a solução apresentada visa não apenas reduzir a exposição a erros clínicos, mas também proporcionar aos profissionais de saúde ferramentas eficientes para enriquecer o seu trabalho diário. Nos subcapítulos subsequentes, delineiam-se os principais contributos decorrentes deste trabalho e apresentam-se propostas para investigações futuras, consolidando assim a relevância e o alcance das inovações introduzidas.

6.1 Principais contributos

Como vimos anteriormente, o projeto delineado ao longo desta dissertação teve como foco central a investigação de uma questão vital que pretendia dar resposta à importância da utilização de uma aplicação de análise de dados no ramo da medicina e na melhoria da qualidade de vida dos pacientes. Esta questão, central ao longo do trabalho, direcionou todos os esforços para compreender os benefícios e as implicações da utilização de tecnologias de análise de dados na prática médica e no bem-estar dos indivíduos atendidos.

Para responder à questão proposta, foi desenvolvida a *AnalyticCare*, uma plataforma de monitorização de análises clínicas, nas instituições de saúde. A plataforma permite traçar tendências por faixa etária e por localidade e, também alerta o profissional de saúde sempre que se registem valores anormais nas análises clínicas dos pacientes. Com isto o propósito desta plataforma é auxiliar as ações diárias dos profissionais de saúde, proporcionando um suporte valioso para a tomada de decisões clínicas. A importância desta plataforma traduziu-se no objetivo da potencial diminuição dos erros clínicos, de médio a longo prazo, durante a interpretação das análises clínicas dos pacientes por parte dos médicos, o que permite aprimorar a prestação de cuidados de saúde.

Para avaliar o impacto da plataforma, foi efetuada uma análise **SWOT**, que permite delinear estratégias futuras e alternativas a opções implementadas, durante o desenvolvimento da *AnalyticCare*, caso se verifiquem impactos menos positivos. O estudo aprofundado desta análise **SWOT** está contemplado no capítulo 5.

6.2 Perspetiva para trabalho futuro

Na procura contínua por aprimorar a plataforma de monitorização de análises clínicas, várias áreas de desenvolvimento futuro podem ser exploradas. Esta secção delineia algumas direções promissoras que poderiam enriquecer ainda mais a aplicação.

Uma abordagem importante para o futuro, envolve partilhar a plataforma *web* de desenvolvimento com profissionais de saúde. Obter *feedback* direto de profissionais de saúde permitirá uma avaliação mais precisa da utilidade da aplicação na prática clínica.

Como parte de futuras melhorias, poderá considerar-se a transição para a utilização de códigos **LOINC** em substituição aos nomes das designações. Os códigos **LOINC** oferecem uma padronização eficiente e globalmente reconhecida para identificar observações laboratoriais, promovendo interoperabilidade e fa-

cilitando comparações entre diferentes sistemas. Explorar benefícios da adoção de códigos **LOINC** pode proporcionar uma estrutura mais robusta e padronizada para a representação das análises clínicas na plataforma. Assim, com as designações a serem interpretadas por códigos, será possível eliminar possíveis ambiguidades associadas a erros ortográficos, simplificando o manuseio e interpretação dos dados. Esta abordagem não apenas aprimora a precisão das análises, como também reduz a complexidade do processamento inicial das informações.

Antecipando um possível aumento significativo de dados, uma consideração importante é a substituição da base de dados SQLite por soluções mais robustas, como Oracle SQL ou outras que possam gerir volumes substanciais de informações. Avaliar a escalabilidade da infraestrutura da base de dados é crucial para garantir o desempenho contínuo da aplicação à medida que o número de utilizadores e dados cresce.

Ao direcionar esforços para estas áreas, a plataforma tem o potencial de se tornar uma ferramenta ainda mais valiosa na monitorização de análises clínicas, contribuindo efetivamente para aprimorar a qualidade dos cuidados de saúde prestados aos pacientes.

Bibliografia

- [1] Eunji Lee. Taming performance variability of healthcare data service frameworks with proactive and coarse-grained memory cleaning. *International journal of environmental research and public health*, 16(17):3096, 2019.
- [2] Vitaly Herasevich, Murat Yilmaz, Hasrat Khan, Rolf D Hubmayr, and Ognjen Gajic. Validation of an electronic surveillance system for acute lung injury. *Intensive care medicine*, 35(6):1018–1023, 2009.
- [3] Marie Egan. Clinical dashboards: impact on workflow, care quality, and patient safety. *Critical care nursing quarterly*, 29(4):354–361, 2006.
- [4] Tomas Hambili Paulo Sanjuluca, Anabela Antunes de Almeida, and Ricardo Cruz-Correia. Assessing the use of hospital information systems (his) to support decision-making: A cross-sectional study in public hospitals in the huila health region of southern angola. In *Healthcare*, volume 10, page 1267. MDPI, 2022.
- [5] Main Naser Alolayyan, Mohammad S Alyahya, Abdallah Hasan Alalawin, Aftab Shoukat, and Farid T Nusairat. Health information technology and hospital performance the role of health information quality in teaching hospitals. *Heliyon*, 6(10), 2020.
- [6] Jahanpour Alipour, Afsaneh Karimi, Saeid Ebrahimi, Fatemeh Ansari, and Yousef Mehdipour. Success or failure of hospital information systems of public hospitals affiliated with zahedan university of medical sciences: A cross sectional study in the southeast of iran. *International journal of medical informatics*, 108:49–54, 2017.
- [7] Jahanpour Alipour, Yousef Mehdipour, and Pezhman Sheibani-Nasab. Evaluation of hospital information systems data quality in educational hospitals of zahedan university of medical sciences form users' perspective. *Journal of Modern Medical Information Sciences*, 5(1):3–11, 2019.

- [8] Mehdi Kahouei and Zainab Abbasi. The prioritization of effective factors on electronic health information security in medical centers. *Journal of Health Information Systems and Informatics*, page 170, 2015.
- [9] Faezeh Abbasi, Reza Khajouei, and Moghaddameh Mirzaee. The efficiency and effectiveness of surgery information systems in iran. *BMC medical informatics and decision making*, 20(1):1–8, 2020.
- [10] Jahanpour Alipour, Yousef Mehdipour, Afsaneh Karimi, Mohadeseh Khorashadizadeh, and Maryam Akbarpour. Security, confidentiality, privacy and patient safety in the hospital information systems from the users' perspective: A cross-sectional study. *International Journal of Medical Informatics*, 175:105066, 2023.
- [11] Maslin Masrom and Ailar Rahimly. Overview of data security issues in hospital information systems. *Pacific Asia Journal of the Association for Information Systems*, 7(4):5, 2015.
- [12] Mehraeen Esmaeel, Ahmadi Maryam, Shajarat Maedeh, and Khoshgam Masumeh. Assessment of hospital information system in selected hospitals in tehran. *Payavard Salamat*, 6(6), 2013.
- [13] Mehmet Saluvan and Al Ozonoff. Functionality of hospital information systems: results from a survey of quality directors at turkish hospitals. *BMC medical informatics and decision making*, 18:1–12, 2018.
- [14] Valéria Alexandra Mendes Garcia. Sistemas de informação em saúde.
- [15] Khilda Nistrina and ATB Bon. Information security for hospital information system using cobit 5 framework. In *Proceedings of the International Conference on Industrial Engineering and Operations Management, Bangkok*, 2019.
- [16] Runtong Zhang, Donghua Chen, Xiaopu Shang, Xiaomin Zhu, and Kecheng Liu. A knowledge-constrained access control model for protecting patient privacy in hospital information systems. *IEEE journal of biomedical and health informatics*, 22(3):904–911, 2017.
- [17] Kiwamu Nagoshi, Takashi Watari, and Yasushi Matsumura. Prospects of hospital information systems and patient safety in japan. *Healthcare Informatics Research*, 28(2):105–111, 2022.
- [18] Maggi Banning. A review of clinical decision making: models and current research. *Journal of clinical nursing*, 17(2):187–195, 2008.

- [19] Doug Laney et al. 3d data management: Controlling data volume, velocity and variety. *META group research note*, 6(70):1, 2001.
- [20] Junghee Jo and Kang-Woo Lee. Mapreduce-based d_elt framework to address the challenges of geospatial big data. *ISPRS International Journal of Geo-Information*, 8(11):475, 2019.
- [21] Karthik Cottur and Veena Gadad. Design and development of data pipelines. *Int. Res. J. Eng. Technol.(IRJET) IOIO*, 7:2715–2718, 2020.
- [22] Athira Nambiar and Divyansh Mundra. An overview of data warehouse and data lake in modern enterprise data management. *Big Data and Cognitive Computing*, 6(4):132, 2022.
- [23] William H Inmon. *Building the data warehouse*. John wiley & sons, 2005.
- [24] Pravin Chandra and Manoj K Gupta. Comprehensive survey on data warehousing research. *International Journal of Information Technology*, 10(2):217–224, 2018.
- [25] Dora Simões. Enterprise data warehouses: A conceptual framework for a successful implementation. In *Canadian Council for Small Business & Entrepreneurship Annual Conference (CCSBE 2010)*. Calgary, Canada, 2010.
- [26] Mutaz M Al-Debei. Data warehouse as a backbone for business intelligence: Issues and challenges. *European Journal of Economics, Finance and Administrative Sciences*, 33(1):153–166, 2011.
- [27] Rihan Hai, Christoph Quix, and Matthias Jarke. Data lake concept and systems: a survey. *arXiv preprint arXiv:2106.09592*, 2021.
- [28] Elisabeta Zagan and Mirela Danubianu. Data lake approaches: A survey. In *2020 International Conference on Development and Application Systems (DAS)*, pages 189–193. IEEE, 2020.
- [29] Mohamed Cherradi and Anass EL Haddadi. Data lakes: A survey paper. In *Innovations in Smart Cities Applications Volume 5: The Proceedings of the 6th International Conference on Smart City Applications*, pages 823–835. Springer, 2022.
- [30] James Dixon. Pentaho, hadoop, and data lakes. *blog*, Oct, 2010.
- [31] Asma Dhaouadi, Khadija Bousselmi, Mohamed Mohsen Gammoudi, Sébastien Monnet, and Slimane Hammoudi. Data warehousing process modeling from classical approaches to new trends: Main features and comparisons. *Data*, 7(8):113, 2022.

- [32] Panos Vassiliadis. Data warehouse modeling and quality issues. *National Technical University of Athens Zographou, Athens, GREECE*, 2000.
- [33] Kamal Kakish and Theresa A Kraft. Etl evolution for real-time data warehousing. In *Proceedings of the Conference on Information Systems Applied Research ISSN*, volume 2167, page 1508, 2012.
- [34] Ralph Kimball, Margy Ross, Warren Thorthwaite, Bob Becker, and Joy Mundy. *The data warehouse lifecycle toolkit*. John Wiley & Sons, 2008.
- [35] Juan Trujillo and Sergio Luján-Mora. A uml based approach for modeling etl processes in data warehouses. In *Conceptual Modeling-ER 2003: 22nd International Conference on Conceptual Modeling, Chicago, IL, USA, October 13-16, 2003. Proceedings 22*, pages 307–320. Springer, 2003.
- [36] Jaiteg Singh. Etl methodologies, limitations and framework for the selection and development of an etl tool. *Int. J. Res. Eng. Appl. Sci*, 6:108–112, 2016.
- [37] Lilia Munoz, Jose-Norberto Mazon, and Juan Trujillo. Systematic review and comparison of modeling etl processes in data warehouse. In *5th Iberian Conference on Information Systems and Technologies*, pages 1–6. IEEE, 2010.
- [38] Gautam Ray, Waleed A Muhanna, and Jay B Barney. Information technology and the performance of the customer service process: A resource-based analysis. *MIS quarterly*, pages 625–652, 2005.
- [39] Enabling Collaborative Care and Fulfilling Healthcare Reform. The promise of mobile technology.
- [40] Neda Tavassoli, Philippe de Souto Barreto, Caroline Berbon, Celine Mathieu, Justine de Kerimel, Christine Lafont, Catherine Takeda, Isabelle Carrie, Antoine Piau, Tania Jouffrey, et al. Implementation of the who integrated care for older people (icope) programme in clinical practice: a prospective study. *The Lancet Healthy Longevity*, 3(6):e394–e404, 2022.
- [41] Michael N Cantor and Lorna Thorpe. Integrating data on social determinants of health into electronic health records. *Health Affairs*, 37(4):585–590, 2018.
- [42] Ou Stella Liang, Jeanne S Sheffield, and Casey Overby Taylor. Detecting patterns of prescription drug use during pregnancy and lactation with visualization techniques. *AMIA Summits on Translational Science Proceedings*, 2019:478, 2019.

- [43] Yichuan Wang, LeeAnn Kung, William Yu Chung Wang, and Casey G Cegielski. An integrated big data analytics-enabled transformation model: Application to health care. *Information & Management*, 55(1):64–79, 2018.
- [44] KA Narayan and M Siva Durga Prasad Nayak. Need for interactive data visualization in public health practice: examples from india. *International Journal of Preventive Medicine*, 12, 2021.
- [45] Venketesh Palanisamy and Ramkumar Thirunavukarasu. Implications of big data analytics in developing healthcare frameworks—a review. *Journal of King Saud University-Computer and Information Sciences*, 31(4):415–425, 2019.
- [46] Alfredo Vellido. The importance of interpretability and visualization in machine learning for applications in medicine and health care. *Neural computing and applications*, 32(24):18069–18083, 2020.
- [47] Christos Chatzakis, Dimitris Floros, Anastasios Liberis, Aggeliki Gerede, Konstantinos Dinas, Nikos Pitsianis, and Alexandros Sotiriadis. Stork: Collaborative online monitoring of pregnancies complicated with gestational diabetes mellitus. In *Healthcare*, volume 10, page 653. MDPI, 2022.
- [48] Miguel Pestana, Ruben Pereira, and Sérgio Moro. Improving health care management in hospitals through a productivity dashboard. *Journal of Medical Systems*, 44:1–19, 2020.
- [49] Katerina A Jones, Frank D Steyn, and Lorraine S Wallace. A longitudinal descriptive assessment of us college covid-19 dashboards. *Journal of American College Health*, pages 1–7, 2022.
- [50] Jack Weiner, Venugopal Balijepally, and Mohan Tanniru. Integrating strategic and operational decision making using data-driven dashboards: the case of st. joseph mercy oakland hospital. *Journal of Healthcare Management*, 60(5):319–330, 2015.
- [51] Aoyu Wu, Yun Wang, Mengyu Zhou, Xinyi He, Haidong Zhang, Huamin Qu, and Dongmei Zhang. Multivision: Designing analytical dashboards with deep learning based recommendation. *IEEE Transactions on Visualization and Computer Graphics*, 28(1):162–172, 2021.
- [52] Sabyasachi Dash, Sushil Kumar Shakyawar, Mohit Sharma, and Sandeep Kaushik. Big data in healthcare: management, analysis and future prospects. *Journal of Big Data*, 6(1):1–25, 2019.
- [53] Su Myat Thaung, Hla Myo Tun, Khin Kyu Kyu Win, Myint Myint Than, and Aye Su Su Phyo. Exploratory data analysis based on remote health care monitoring system by using iot. *Communications*, 8(1):1–8, 2020.

- [54] A Lohachab. Research anthology on big data analytics, architectures, and applications.
- [55] Meghan Reading Turchioe, Annie Myers, Samuel Isaac, Dawon Baik, Lisa V Grossman, Jessica S Ancker, and Ruth Masterson Creber. A systematic review of patient-facing visualizations of personal health data. *Applied clinical informatics*, 10(04):751–770, 2019.
- [56] SA Senthilkumar, Bharatendara K Rai, Amruta A Meshram, Angappa Gunasekaran, and S Chandrakumarmangalam. Big data in healthcare management: a review of literature. *American Journal of Theoretical and Applied Business*, 4(2):57–69, 2018.
- [57] Joana Muchagata and Ana Ferreira. How can visualization affect security? In *ICEIS (2)*, pages 503–510, 2018.
- [58] Cornelia Gyöřödi, Robert Gyöřödi, Alexandra Țtefan, and Livia Bandici. A comparative study of databases with different methods of internal data management. *Database*, 7(4), 2016.
- [59] Hui Lv, Yonglin Yu, Tan Shu, Dexiu Huang, Shan Jiang, and Liam P Barry. Photonic generation of ultra-wideband signals by direct current modulation on soa section of an soa-integrated sgdb laser. *Optics Express*, 18(7):7219–7227, 2010.
- [60] Edgar F Codd. A relational model of data for large shared data banks. *Communications of the ACM*, 13(6):377–387, 1970.
- [61] Bárbara Martins, Diana Ferreira, Cristiana Neto, António Abelha, and José Machado. Data mining for cardiovascular disease prediction. *Journal of Medical Systems*, 45:1–8, 2021.
- [62] Mohammad A Hassan. Relational and nosql databases: The appropriate database model choice. In *2021 22nd International Arab Conference on Information Technology (ACIT)*, pages 1–6. IEEE, 2021.
- [63] Jing Han, Meina Song, and Junde Song. A novel solution of distributed memory nosql database for cloud computing. In *2011 10th IEEE/ACIS International Conference on Computer and Information Science*, pages 351–355. IEEE, 2011.
- [64] Bogdan George Tudorica and Cristian Bucur. A comparison between several nosql databases with comments and notes. In *2011 RoEduNet international conference 10th edition: Networking in education and research*, pages 1–5. IEEE, 2011.

- [65] Ahsan Malik, Aqil Burney, and Fawad Ahmed. A comparative study of unstructured data with sql and no-sql database management systems. *Journal of Computer and Communications*, 8(4):59–71, 2020.
- [66] Dam Quang Tuan, Seungyong Cheon, and Youjip Won. On the io characteristics of the sqlite transactions. In *Proceedings of the International Conference on Mobile Software Engineering and Systems*, pages 214–224, 2016.
- [67] Hyojun Kim, Nitin Agrawal, and Cristian Ungureanu. Revisiting storage for smartphones. *ACM Transactions on Storage (TOS)*, 8(4):1–25, 2012.
- [68] Yaoming Chen, Wei Shen, Hongwei Huo, and Youzhi Xu. A smart gateway for health care system using wireless sensor network. In *2010 Fourth International Conference on Sensor Technologies and Applications*, pages 545–550. IEEE, 2010.
- [69] Charalampos Doukas, Thomas Pliakas, and Ilias Maglogiannis. Mobile healthcare information management utilizing cloud computing and android os. In *2010 Annual International Conference of the IEEE Engineering in Medicine and Biology*, pages 1037–1040. IEEE, 2010.
- [70] Vikas Vik Kumar. *Healthcare analytics made simple: techniques in healthcare computing using machine learning and Python*. Packt Publishing Ltd, 2018.
- [71] Leire Lopez-Samaniego and Begonya Garcia-Zapirain. A robot-based tool for physical and cognitive rehabilitation of elderly people using biofeedback. *International journal of environmental research and public health*, 13(12):1176, 2016.
- [72] Paschou Mersini, Evangelos Sakkopoulos, and Athanasios Tsakalidis. Appification of hospital healthcare and data management using qr codes. In *IISA 2013*, pages 1–6. IEEE, 2013.
- [73] Joontaek Oh and Youjip Won. Embedded dbms design for in-vehicle information management. In *2018 IEEE 7th Non-Volatile Memory Systems and Applications Symposium (NVMSA)*, pages 111–112. IEEE, 2018.
- [74] Rakesh Kumar, Neha Gupta, Shilpi Charu, Kanishk Jain, and Sunil Kumar Jangir. Open source solution for cloud computing platform using openstack. *International Journal of Computer Science and Mobile Computing*, 3(5):89–98, 2014.

- [75] Sam M Ireland and Andrew CR Martin. Graphql for the delivery of bioinformatics web apis and application to zincbind. *Bioinformatics Advances*, 1(1):vbab023, 2021.
- [76] Lincoln Stein. Creating a bioinformatics nation. *Nature*, 417(6885):119–120, 2002.
- [77] Pieter BT Neerincx and Jack AM Leunissen. Evolution of web services in bioinformatics. *Briefings in bioinformatics*, 6(2):178–188, 2005.
- [78] Cristiana Neto, Maria Brito, Vitor Lopes, Hugo Peixoto, António Abelha, and José Machado. Application of data mining for the prediction of mortality and occurrence of complications for gastric cancer patients. *Entropy*, 21(12):1163, 2019.
- [79] Regina Sousa, Rui Miranda, Ailton Moreira, Carlos Alves, Nicolas Lori, and José Machado. Software tools for conducting real-time information processing and visualization in industry: An up-to-date review. *Applied Sciences*, 11(11):4800, 2021.
- [80] Roy Thomas Fielding. *Architectural styles and the design of network-based software architectures*. University of California, Irvine, 2000.
- [81] Cristiana Neto, Fábio Senra, Jaime Leite, Nuno Rei, Rui Rodrigues, Diana Ferreira, and José Machado. Different scenarios for the prediction of hospital readmission of diabetic patients. *Journal of Medical Systems*, 45:1–9, 2021.
- [82] Kai Lei, Yining Ma, and Zhi Tan. Performance comparison and evaluation of web development technologies in php, python, and node. js. In *2014 IEEE 17th international conference on computational science and engineering*, pages 661–668. IEEE, 2014.
- [83] Lionel Nkenyereye and Jong-Wook Jang. Performance evaluation of server-side javascript for healthcare hub server in remote healthcare monitoring system. *Procedia Computer Science*, 98:382–387, 2016.
- [84] Erik Wittern, Philippe Suter, and Shriram Rajagopalan. A look at the dynamics of the javascript package ecosystem. In *Proceedings of the 13th International Conference on Mining Software Repositories*, pages 351–361, 2016.
- [85] Ahmad Abdellatif, Yi Zeng, Mohamed Elshafei, Emad Shihab, and Weiyi Shang. Simplifying the search of npm packages. *Information and Software Technology*, 126:106365, 2020.
- [86] Joel Vainikka. Full-stack web development using django rest framework and react. 2018.

- [87] Carlos Santana Roldán. *React Design Patterns and Best Practices: Design, build and deploy production-ready web applications using standard industry practices*. Packt Publishing Ltd, 2019.
- [88] Sanchit Aggarwal et al. Modern web-development using reactjs. *International Journal of Recent Research Aspects*, 5(1):133–137, 2018.
- [89] Marisa Esteves, Márcia Esteves, António Abelha, and José Machado. A proof of concept of a business intelligence platform to support the decision-making process of health professionals in waiting lists. In *Research Anthology on Decision Support Systems and Decision Management in Healthcare, Business, and Engineering*, pages 1015–1034. IGI Global, 2021.
- [90] Radha Balamuralikrishna and John C Dugger. Swot analysis—a management tool for initiating new programs in vocational schools. *Journal of Vocational and Technical Education*, 12(1):36–41, 1995.
- [91] Otto H Chang and Chee W Chow. The balanced scorecard: A potential tool for supporting change and continuous improvement in accounting education. *Issues in Accounting Education*, 14(3):395–412, 1999.
- [92] Thomas J Chermack and Bernadette K Kasshanna. The use and misuse of swot analysis and implications for hrd professionals. *Human Resource Development International*, 10(4):383–399, 2007.
- [93] Dag Øivind Madsen. Swot analysis: a management fashion perspective. *International Journal of Business Research*, 16(1):39–56, 2016.
- [94] T Hill and R Westbrook. Swot analysis: It's time for a product recall, long range planning, 30 (1): 46-52 [https://doi.org/10.1016/S0024-6301\(96\)00095-7](https://doi.org/10.1016/S0024-6301(96)00095-7), pages 00095–7, 1997.
- [95] Rod Kuhn King. Enhancing swot analysis using triz and the bipolar conflict graph: a case study on the microsoft corporation. *Proceedings of TRIZCON2004, 6th Annual Altshuller Institute*, pages 25–27, 2004.
- [96] George Panagiotou. Bringing swot into focus. *Business strategy review*, 14(2):8–10, 2003.
- [97] Ali Hadighi and Iraj Mahdavi. A new model for strategy formulation using mahalanobis-taguchi system and clustering algorithm. 2011.
- [98] Birger Wernerfelt. A resource-based view of the firm. *Strategic management journal*, 5(2):171–180, 1984.

- [99] Robert E Hoskisson, William P Wan, Daphne Yiu, and Michael A Hitt. Theory and research in strategic management: Swings of a pendulum. *Journal of management*, 25(3):417–456, 1999.
- [100] Claudia Brito, Marisa Esteves, Hugo Peixoto, António Abelha, and José Machado. A data mining approach to classify serum creatinine values in patients undergoing continuous ambulatory peritoneal dialysis. *Wireless Networks*, pages 1–9, 2022.
- [101] Dac Teoli, Terrence Sanvictores, and Jason An. Swot analysis. 2019.