



UNIVERSIDADE ESTADUAL PAULISTA  
"JÚLIO DE MESQUITA FILHO"  
Câmpus de São José do Rio Preto



DOUGLAS BRANDÃO DOS SANTOS

## **Visualização de dados estruturados e não estruturados da área da saúde**

São José do Rio Preto  
2021

DOUGLAS BRANDÃO DOS SANTOS

## **Visualização de dados estruturados e não estruturados da área da saúde**

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Orientador: Prof. Dr. Carlos Roberto Valêncio

São José do Rio Preto  
2021

S237v Santos, Douglas Brandão dos  
Visualização de dados estruturados e não estruturados da área da saúde / Douglas Brandão dos Santos. -- São José do Rio Preto, 2021  
74 p. : il., tabs.

Trabalho de conclusão de curso (Bacharelado - Ciência da Computação) - Universidade Estadual Paulista (Unesp), Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto  
Orientador: Carlos Roberto Valêncio

1. Extração de conhecimento. 2. Visual Data Mining. 3. Visualização de dados. 4. Computação em nuvem. 5. Banco de dados.  
I. Título.

Sistema de geração automática de fichas catalográficas da Unesp. Biblioteca do Instituto de Biociências Letras e Ciências Exatas, São José do Rio Preto. Dados fornecidos pelo autor(a).

Essa ficha não pode ser modificada.

DOUGLAS BRANDÃO DOS SANTOS

**Visualização de dados estruturados e não estruturados  
da área da saúde**

Trabalho de Conclusão de Curso apresentado ao Departamento de Ciências de Computação e Estatística do Instituto de Biociências Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, como parte dos requisitos necessários para obtenção do grau de Bacharel em Ciência da Computação.

Banca examinadora:

Prof. Dr. Carlos Roberto Valêncio  
UNESP – São José do Rio Preto  
Orientador

Prof. Dr. Geraldo Francisco Donega Zafalon  
UNESP – São José do Rio Preto

Prof<sup>a</sup>. Dr<sup>a</sup>. Carina Alexandra Rondini  
UNESP – São José do Rio Preto

São José do Rio Preto  
2021

# Agradecimentos

Agradeço inicialmente a Deus, por estar ao meu lado ao longo de todo o caminho necessário para chegar até aqui, me dando suporte quando mais precisei.

À minha mãe, Lucia, e meu irmão, Jardel, pelo apoio incondicional durante todo meu trajeto acadêmico. Sem vocês, eu não teria qualquer razão para chegar onde cheguei, portanto, obrigado por estarem sempre ao meu lado.

A meu amado pai, que sempre me incentivou e me inspirou. Espero que o senhor esteja vendo tudo de onde está e saiba que eu não chegaria aqui sem seus ensinamentos. Sou extremamente grato e devo tudo que sei a você, te amo e sempre te amarei.

À minha amada Beatriz, que esteve sempre ao meu lado e acreditou em mim, em momentos que nem mesmo eu acreditei.

Aos meus amigos da faculdade e do Grupo de Banco de Dados, em especial ao Prof. Valêncio. Obrigado por terem me acolhido e tornado esse caminho mais leve. Espero que, independentemente de nossos destinos, sempre possamos nos reunir.

Todos vocês tiveram fundamental importância nessa trajetória e são absolutamente essenciais na minha vida. Nenhum tipo de sucesso seria possível sem vocês, por isso, agradeço a todos do fundo do meu coração.

*“Não há doença no mundo que não possa ser curada!”*

Tony Tony Chopper

# Resumo

A maior adoção de tecnologia da informação no âmbito de assistência médica em conjunto com a maior disponibilidade de dados observada nos últimos anos apresenta o desafio de utilizar técnicas de análise de dados com o intuito de proporcionar benefícios, como tomadas de decisões mais assertivas e melhorias na qualidade dos serviços prestados pela área. O processo de encontrar padrões úteis meio a grandes volumes de dados é definido como *data mining* e pode ser implementado por meio de diversas técnicas, dentre elas, a visualização de dados. Para que o processo de *data mining* seja introduzido na extração de conhecimento de forma mais efetiva, é de suma importância a inclusão do fator humano durante a exploração de dados. Dessa forma, combina-se a flexibilidade, criatividade e conhecimentos gerais do ser humano com a grande capacidade de processamento e de armazenamento proveniente dos computadores atuais. Neste cenário, as técnicas de *data mining* baseadas em visualização surgem com o intuito de integrar o fator humano à mineração de dados, de forma que este possa aplicar suas habilidades perceptivas ao conjunto de dados analisados. Porém, nota-se que as aplicações de visualização de dados presentes na literatura científica falham em lidar com a falta de estruturação dos dados provenientes da área da saúde, característica presente em uma parcela significativa dos dados deste domínio, além de não suportarem os requisitos de escalabilidade do mesmo. Portanto, o presente trabalho propõe uma arquitetura flexível baseada em microserviços eficiente em lidar com as diferentes características dos dados presentes no setor da saúde. Além disso, foi desenvolvido um ambiente de visualização de dados com base na arquitetura proposta, de modo que este foi avaliado por meio de uma abordagem heurística na qual obteve avaliação positiva dos usuários e alcançou um percentual de 75,80% da pontuação máxima prevista no método de validação do sistema.

**Palavras-chave:** Banco de dados, *Visual Data Mining*, visualização em banco de dados.

# Abstract

The greater adoption of information technology in the field of medical care, together with the greater availability of data observed in recent years, presents the challenge of using data analysis techniques in order to provide benefits, such as more assertive decision-making and improvements in quality of services provided by the area. The process of finding useful patterns through large volumes of data is defined as data mining and can be implemented through several techniques, including data visualization. In order for the data mining process to be introduced in the extraction of knowledge more effectively, it is of paramount importance to include the human factor during data exploration. In this way, the flexibility, creativity and general knowledge of human beings are combined with the great processing and storage capacity of today's computers. In this scenario, visualization-based data mining techniques emerge with the aim of integrating the human factor into data mining, so that the latter can apply their perceptual skills to the analyzed data set. However, it is noted that the data visualization applications present in the scientific literature fail to deal with the lack of structuring of data from the health area, a characteristic present in a significant portion of the data in this domain, in addition to not supporting the requirements of scalability of the same. Therefore, the present work proposes a flexible architecture based on microservices that is efficient in dealing with the different characteristics of the data present in the health sector. In addition, a data visualization environment was developed based on the proposed architecture, so that it was evaluated through a heuristic approach in which it obtained a positive evaluation from users and reached a percentage of 75.80% of the maximum score predicted in the system validation method.



# Lista de Ilustrações

Figura 1 – Seis Vs de big data contextualizados para ambientes da saúde .....	23
Figura 2 – Arquitetura do sistema .....	35
Figura 3 – Estrutura de um JSON Web Token .....	37
Figura 4 – Fluxo de execução dos microsserviços .....	39
Figura 5 – Tela de autenticação do ambiente de visualização.....	40
Figura 6 – Tela inicial do ambiente de visualização .....	42
Figura 7 – Gráfico de consultas médicas presente na página Home .....	42
Figura 8 - Gráfico de emergências presente na página Home .....	43
Figura 9 - Gráfico de Internações presente na página Home .....	43
Figura 10 - Gráfico de Altas Médicas presente na página Home .....	44
Figura 11 - Recursos de manipulação.....	44
Figura 12 - Tela de pacientes.....	45
Figura 13 - Gráfico de Visão geral de pacientes presente na página Pacientes.....	46
Figura 14 - Gráfico de Pacientes hospitalizados presente na página de Pacientes.....	46
Figura 15 - Página de Internações .....	47
Figura 16 - Gráfico de CIDs da Internação presente na página de Internação .....	48
Figura 17 - Tela de atendimentos de emergência de um paciente.....	49
Figura 18 - Gráfico dde CIDs de Emergência presente na página de Emergências.....	49
Figura 19 – Resultados da avaliação dos princípios heurísticos .....	56

# Lista de Tabelas

Tabela 1 – Resultados do teste de uso .....	53
Tabela 2 – Resultados do fator Controle e liberdade do usuário.....	54
Tabela 3 – Resultados do fator Flexibilidade e eficiência de uso .....	55
Tabela 4 – Resultados do fator Orientação e ajuda .....	55

# Lista de Quadros

Quadro 1 - Comparativo dos trabalhos correlatos .....	33
Quadro 2 – Interfaces disponibilizadas pela aplicação.....	36
Quadro 3 - Comparação com trabalhos correlatos e este trabalho .....	59

# Lista de Abreviaturas e Siglas

API – *Application Programming Interface*

CFM – Conselho Regional de Medicina

CID10 – Classificação Internacional de Doenças, décima versão

COPD – *Chronic Obstructive Pulmonary Disease*

EBPH – *Evidence-Based Public Health*

ELK – *Logstash, Elasticsearch* e Kibana

GBD – Grupo de Banco de Dados

HMAC – *Hash-based Message Authentication Code*

HTTP – *Hypertext Transfer Protocol*

IBILCE – Instituto de Biociências, Letras e Ciências Exatas

JSON – *JavaScript Object Notation*

JWT – *JSON Web Token*

KDD – *Knowledge Discovery in Databases*

LSTM – *Long-short Term Memory*

PEP – Prontuário Eletrônico do Paciente

PopHR – *Population Health Record*

SPA – *Single Page Application*,

TI – Tecnologia da Informação

UMLS – *Unified Medical Language System*

XML – *eXtensible Markup Language*

# Sumário

<b>Capítulo 1 Introdução .....</b>	<b>14</b>
1.1 Motivação e escopo .....	15
1.2 Objetivo .....	17
1.3 Metodologia.....	17
1.4 Organização da Monografia .....	18
<b>Capítulo 2 Fundamentação teórica .....</b>	<b>19</b>
2.1 Considerações iniciais .....	19
2.2 Prontuário eletrônico do paciente.....	19
2.3 <i>Big Data</i> .....	21
2.4 Visual Data Mining .....	23
2.5 Classificação de técnicas de Visual Data Mining .....	24
2.5.1 Dados a serem visualizados.....	24
2.5.2 Técnicas de visualização .....	25
2.5.3 Técnicas de interação e distorção.....	26
2.6 Computação em nuvem.....	26
2.7 Microsserviços.....	27
2.8 Trabalhos correlatos .....	28
2.8.1 Ambiente de visualização de dados heterogêneos da área da saúde .....	28
2.8.2 Ambiente de visualização de condições médicas extraídas de notas clínicas.....	29
2.8.3 Ambiente de visualização do progresso da doença de obstrução pulmonar crônica..	30
2.8.4 Ambiente de visualização de dados do paciente em forma de dashboard.....	31
2.9 Principais características dos trabalhos correlatos .....	32
2.10 Considerações finais.....	33
<b>Capítulo 3 Desenvolvimento .....</b>	<b>34</b>
3.1 Considerações iniciais .....	34
3.2 Arquitetura do sistema.....	34
3.3 Microsserviço controlador.....	35
3.3.1 Autenticação.....	36
3.3.2 Sistema de mensagens .....	37

3.4	Microserviços de dados hospitalares .....	37
3.4.1	Microserviços de dados estruturados .....	38
3.4.2	Microserviços de dados não estruturados .....	38
3.5	Ambiente de visualização.....	39
3.5.1	Página de autenticação .....	40
3.5.2	Página Home .....	40
3.5.3	Pacientes .....	45
3.5.4	Internações do paciente .....	47
3.5.5	Atendimentos de emergência do paciente .....	48
3.6	Considerações finais.....	50
<b>Capítulo 4 Avaliação experimental .....</b>		<b>51</b>
4.1	Considerações iniciais .....	51
4.2	Metodologia de experimentação .....	51
4.3	Resultados .....	53
4.4	Considerações finais.....	56
<b>Capítulo 5 Conclusões .....</b>		<b>57</b>
5.1	Considerações iniciais .....	57
5.2	Contribuições científicas .....	58
5.3	Trabalhos futuros.....	59
5.4	Considerações finais.....	60

# Capítulo 1

## Introdução

Com os recentes avanços tecnológicos, diversos setores da sociedade moderna expandiram seus investimentos em tecnologia da informação, de modo a melhorar a qualidade de seus serviços. No setor da saúde, tal aumento de investimento possibilitou o surgimento do Prontuário Eletrônico do Paciente (PEP), ferramenta que permite o compartilhamento em tempo real de informações dentro e por meio da interface de atendimento médico (MENACHEMI; COLLUM, 2011). Deste modo, intervenções digitais dentro do PEP, como prescrições eletrônicas ou registros computadorizados de pedidos médicos podem diminuir as chances de ocorrência de erros, de modo a resultar em benefícios como níveis melhores de segurança para os pacientes e potencial economia de recursos financeiros (REDDY, 2015).

Para que os recursos oferecidos pelo PEP sejam utilizados de forma eficaz, as organizações e instituições do setor devem atentar-se a fatores como a capacitação de usuários, que devem possuir conhecimentos a respeito dos dados presentes no sistema e utilização de métodos eficientes de *data mining*. Tal processo tem o objetivo de descobrir estruturas e padrões a princípio camuflados meio ao grande volume de dados gerados por estes sistemas. Além disso, o *data mining* faz parte de um processo mais amplo, responsável pela extração de conhecimentos de alto nível potencialmente úteis, conhecido como descoberta de conhecimento em banco de dados – do inglês, *Knowledge Discovery in Databases* (KDD). Porém, para que o processo de *data mining* seja efetivo, é de suma importância a integração do fator humano durante a exploração de dados, de modo a combinar, a flexibilidade, criatividade e conhecimento geral do humano com a grande capacidade de armazenamento e de

processamento do computador. Neste cenário, técnicas visuais de mineração de dados surgem com o objetivo de integrar os humanos ao processo de exploração de dados. Para isso, os dados a serem explorados são apresentados de maneira visual, normalmente por meio de gráficos que suportam interações com o usuário. Dessa forma, a exploração de dados por intermédio de técnicas de visualização resulta em análises mais rápidas e normalmente, tomadas de decisões mais assertivas.

## 1.1 Motivação e escopo

Ambientes de visualização de dados apresentam-se cada vez mais como uma ferramenta de relevância, uma vez que estes oferecem benefícios para seus usuários, como maior facilidade para ler e compreender os dados contidos nos prontuários eletrônicos, visto que estes são apresentados de forma intuitiva, especialmente em ambientes com características *Big Data* (KO; CHANG, 2017). Na área da saúde, por meio de infográficos que visam reduzir a complexidade da tarefa de processar dados a princípio confusos e díspares, tais ambientes podem oferecer recursos como melhoria no processo de acompanhamento de doenças, maior praticidade e agilidade durante o processo de avaliação de resultados de exames médicos, visualização estatística por meio do uso de tendências de valores de laboratório, de forma a reduzir os esforços necessários para a identificação de *outliers* e possíveis correções de dados, entre outros benefícios (YANDRAPALLI; JONES; PURKAYASTHA, 2019).

Yandrapalli, Jones e Purkayastha (2019), por exemplo, implementaram uma *dashboard* com o intuito de apresentar medidas laboratoriais relacionadas à diabetes de forma centralizada. Dessa forma, no momento em que um médico acessa o registro de determinado paciente, medidas clínicas associadas à diabetes são exibidas por meio de gráficos, de forma intuitiva e personalizável. Por sua vez, Filho et al. (2020) desenvolveram e implementaram uma *dashboard* para o sistema de gerenciamento de prontuários eletrônicos que integra clínicas administradas por alunos da Universidade Federal do Rio Grande do Norte. Por meio da ferramenta construída, é possível realizar análises visuais das informações de pacientes, com o auxílio de gráficos interativos. Assim, o processo de interpretação de dados clínicos e posterior tomada de decisão foi consideravelmente aprimorado.

Apesar de promissores, no que diz respeito ao domínio de saúde, um dos maiores desafios para os trabalhos de visualização de dados encontrados na literatura atual é a grande variedade de dados disponíveis, como por exemplo, dados não estruturados, presentes em estruturas como as notas clínicas (GOTZ; BORLAND, 2016). Estas estruturas podem conter



informações valiosas sobre determinado paciente, como, sintomas, diagnósticos, tratamentos, uso de medicamentos, entre outros (DALIANIS, 2018). Atualmente, dados não estruturados constituem a maior parcela dos dados presentes em prontuários eletrônicos (BETHANY; PERCHA, 2021). Porém, há uma grande dificuldade em se lidar com os dados provenientes das notas clínicas, visto que estas são escritas por diversos profissionais de saúde, como enfermeiros, médicos, fisioterapeutas, psicólogos, entre outros. Isto, aliado ao fato de que normalmente os profissionais escrevem as notas clínicas sob pressão (DALIANIS, 2018), resulta em problemas como, falta de padronização, abreviaturas, erros de digitação, jargões, frases incompletas, entre outros tipos de ruídos que tornam o processo de análise de dados por meio das notas clínicas mais difícil de ser aplicado. (YU et al., 2019).

Outro ponto de atenção a respeito dos ambientes visualização presentes na literatura acadêmica atual refere-se ao desenho de arquitetura utilizada por tais sistemas, bem como a forma como são implantados. Recomenda-se que sistemas pertencentes ao escopo médico devem empregar práticas de reuso de software (DA SILVA et al, 2019). Porém, em sua maioria, os sistemas da área são desenhados com base em arquiteturas que não suportam diferentes componentes em execução de forma independente, fator que limita a reusabilidade dos mesmos (DA SILVA et al, 2019). Além disso, arquiteturas mais flexíveis permitem maior poder de escalabilidade para os sistemas, além de menores custos de implantação e maiores níveis de confiabilidade e disponibilidade (LARRUCEA et al, 2018).

Diante do cenário apresentado, pode-se observar os potenciais benefícios de ambientes de visualização aplicados em dados provenientes de prontuários eletrônicos, uma vez que estes podem ser utilizados em conjunto com o fator humano presente em hospitais e clínicas de modo a aumentar a assertividade da análise e conseqüentemente, das tomadas de decisões. Com tudo, evidencia-se também a falta de ambientes de visualização que trabalhem de forma eficaz os diferentes tipos de dados presentes no domínio da saúde ao mesmo tempo que são capazes de suportar as necessidades de crescimento em larga escala provenientes do setor, de forma a manter a qualidade dos serviços. Dessa forma, o desenvolvimento de ambientes de visualização aprimorados e capazes de suportar os diferentes tipos de dados presentes na área da saúde mostra-se um contribuição científica promissora e de suma importância, uma vez que a análise dos complexos e volumosos conjuntos de dados não estruturados presentes na área pode se tornar uma atividade difícil de ser executada (KAYSI; KESLER, 2018) e técnicas de visualização podem guiar os profissionais de saúde a uma melhor compreensão a respeito dos dados armazenados.

## 1.2 Objetivo

O presente trabalho tem como objetivo a elaboração e proposição de uma arquitetura flexível baseada em microsserviços capaz de lidar com os diferentes tipos de dados presentes no setor da saúde, isto é, dados estruturados e dados não estruturados. Além disso, um ambiente de visualização de dados da saúde baseado na arquitetura proposta foi desenvolvido e validado. Desta forma, este trabalho contribui cientificamente preenchendo uma lacuna identificada na literatura científica, de modo que apresenta uma alternativa para a construção de ambientes de visualização capazes de lidar com os problemas de heterogeneidade dos dados e de escalabilidade apresentados.

## 1.3 Metodologia

A metodologia em que se baseia o desenvolvimento deste trabalho é firmada nas seguintes etapas: levantamento dos conceitos fundamentais e respectivo estado da arte; elaboração dos recursos de *Visual Data Mining*; elaboração e implementação da arquitetura; desenvolvimento do ambiente de visualização; realização das etapas de experimentação.

Inicialmente, um levantamento bibliográfico foi realizado referente a prontuário do paciente, prontuário eletrônico do paciente, *Visual Data Mining*, técnicas de visualização e *Cloud Computing*. Além disso, foi executado um estudo sobre o estado da arte das abordagens de visualização e *Visual Data Mining* relacionados a dados clínicos.

Após as etapas anteriores, iniciou-se o processo de elaboração do ambiente de visualização, onde a arquitetura do sistema foi planejada e foi efetuada uma análise prévia à seleção de atributos e suas respectivas formas de visualização. Em seguida, iniciou-se o desenvolvimento do ambiente de visualização com o auxílio de uma biblioteca de manipulação de elementos gráficos, entre outras tecnologias.

Por fim, após a conclusão do processo de implementação foram realizadas as etapas de experimentação do sistema, que envolvem avaliá-lo segundo critérios de qualidade importantes para os sistemas de análise visual.

## 1.4 Organização da Monografia

Nesta seção foi apresentado o problema de visualização com suporte a grande variedade de dados presente no setor da saúde. Além disso, foi exposta a motivação e objetivo do trabalho proposto, bem como os recursos disponíveis para concretizá-lo. As seções a seguir estão organizadas da seguinte forma:

- a) Seção 2 – Fundamentação teórica: Revisão de conceitos relevantes para o trabalho, como, prontuário do paciente, prontuário eletrônico do paciente, *big data*, técnicas de visualização e *Visual Data Mining*, *Cloud Computing* e microsserviços.
- b) Seção 3 – Ambiente de visualização de dados: Descrição da arquitetura do sistema e do processo de desenvolvimento do ambiente de visualização.
- c) Seção 4 – Validação do projeto: Apresentação do método utilizado para a validação do ambiente, isto é, a forma como as características foram analisadas, bem como os resultados desta análise.
- d) Seção 5 – Conclusão: Apresentação da contribuição científica e sugestões de trabalhos futuros.

# Capítulo 2

## Fundamentação teórica

### 2.1 Considerações iniciais

Neste capítulo será apresentada a fundamentação teórica e o estado da arte por meio dos trabalhos correlatos encontrados na literatura científica. Inicialmente, são abordados os principais conceitos a respeito do prontuário eletrônico do paciente, o processo de *Visual Data Mining* e as principais técnicas de visualização presentes na literatura. Em seguida, é tratado o tema de Computação em Nuvem, como os principais trabalhos acadêmicos definem seus termos e conceitos e porque estes são importantes na área médica. Por fim, os trabalhos correlatos serão apresentados, analisados e comparados entre si.

### 2.2 Prontuário eletrônico do paciente

O Conselho Regional de Medicina (CFM), define no Artigo 1º da Resolução nº 1.638/2002 o conceito de Prontuário Médico do Paciente como documento único baseado em uma coleção de dados gerados a partir de acontecimentos e situações relacionadas com a saúde do paciente e a assistência prestada, sendo estes, de caráter legal, sigiloso e científico, que possibilitam a comunicação interna entre os membros da equipe multiprofissional de saúde e a continuidade da assistência prestada ao indivíduo (BRASIL, 2002). Desta forma, pode-se entender o Prontuário Eletrônico do Paciente (PEP) como um registro eletrônico relativo a informações de saúde de determinado paciente, geradas por um ou mais encontros em qualquer ambiente médico. Tais informações incluem, mas não se limitam a, dados demográficos do paciente, notas para o acompanhamento do progresso do paciente, medicamentos utilizados,

histórico médico anterior, histórico de imunizações, dados referentes a exames laboratoriais, relatórios de radiologia, entre outros (MENACHEMI; COLLUM, 2011).

A utilização do PEP apresenta-se como uma estratégia promissora para melhorar a qualidade dos diferentes serviços prestados ao paciente e aumentar a produtividade e eficiência nos âmbitos clínicos (REDDY, 2015). Além disso, pode-se destacar ainda procedimentos em particular que são significativamente aprimorados com a adoção do PEP, como: suporte à decisão clínica, automatização de entradas de pedidos médicos e intercâmbio eletrônico de informações de saúde, que poderão contar com ferramentas aperfeiçoadas (MENACHEMI; COLLUM, 2011). Entre outros benefícios, pode-se observar ainda: impacto nos custos relativos ao modo tradicional de operação das entidades de saúde, como, redução de custos em suprimentos de papelaria, melhor utilização de testes, redução dos custos de escrita para processos administrativos manuais e maior satisfação dos profissionais da área médica, visto que a menor utilização de documentos físicos pode reduzir o tempo de diversas tarefas (REDDY, 2015).

Contudo, foram identificados desafios em utilizar os prontuários eletrônicos do paciente, muitas vezes relacionados com à qualidade e validação dos dados, e heterogeneidade entre diferentes sistemas (COWIE et al., 2017). No que diz respeito aos dados armazenados no PEP, estes podem possuir diferentes formatos, como gráficos, símbolos, texto livre, números, entre outros. Dessa forma, torna-se conveniente classificá-los em dois grandes grupos: dados estruturados e não estruturados (TAYEFI et al., 2021). Pode-se mencionar como exemplos de dados estruturados informações como idade e gênero do paciente, peso, altura, pressão sanguínea, resultados de testes laboratoriais, medicações ministradas, entre outros. Por outro lado, dados não estruturados apresentam-se como notas clínicas, registros cirúrgicos, resumos de alta, relatórios de processos radiológicos, imagens e relatórios de patologia que são armazenados nos prontuários eletrônicos (SUN et al., 2018).

Os dados estruturados presentes no prontuário podem ser processados de forma relativamente fácil por meio dos métodos presentes na literatura atual. Porém, apenas os dados estruturados não oferecem informações suficientes para a construção de um contexto clínico completo de determinado paciente (TAYEFI et al., 2021).

Em contraste, os dados não estruturados podem fornecer informações extras e valiosas sobre o paciente, porém, analisá-los exige maior complexidade de processamento, o que consome mais tempo e requer esforço manual, dado que estes podem apresentar diversos obstáculos para realização da análise, como erros gramaticais, erros de escrita, linguagem e

termos locais, ambiguidade semântica, siglas não padronizadas, entre outros (SUN et al., 2018). Portanto, apresenta-se como desafio realizar o processamento automatizado de dados estruturados e não estruturados, visto que uma ferramenta de suporte à tomada de decisão bem planejada deve ser eficiente em utilizar ambos os formatos de dados, a fim de extrair informações úteis e valiosas que podem oferecer benefícios para o paciente (TAYEFI et al., 2021).

Além disso, os sistemas de gerenciamento de prontuário eletrônico apresentam alta heterogeneidade, conforme citado anteriormente, desta forma, problemas de interoperabilidade são comuns, uma vez que a maior parte dos softwares de gerenciamento de PEP não são construídos para oferecer suporte à comunicação com aplicações externas e o desenvolvimento de interfaces específicas pode ser custoso (REDDY, 2015). Os softwares de gerenciamento de prontuário eletrônico devem possuir uma arquitetura flexível, de forma a adaptar-se às novas intervenções e requisitos (COWIE et al., 2017), porém, pesquisadores apontam que grande parte dos softwares presentes nas aplicações da área da saúde não apresentam arquiteturas que forneçam tal flexibilidade (DA SILVA et al., 2019).

### **2.3 *Big Data***

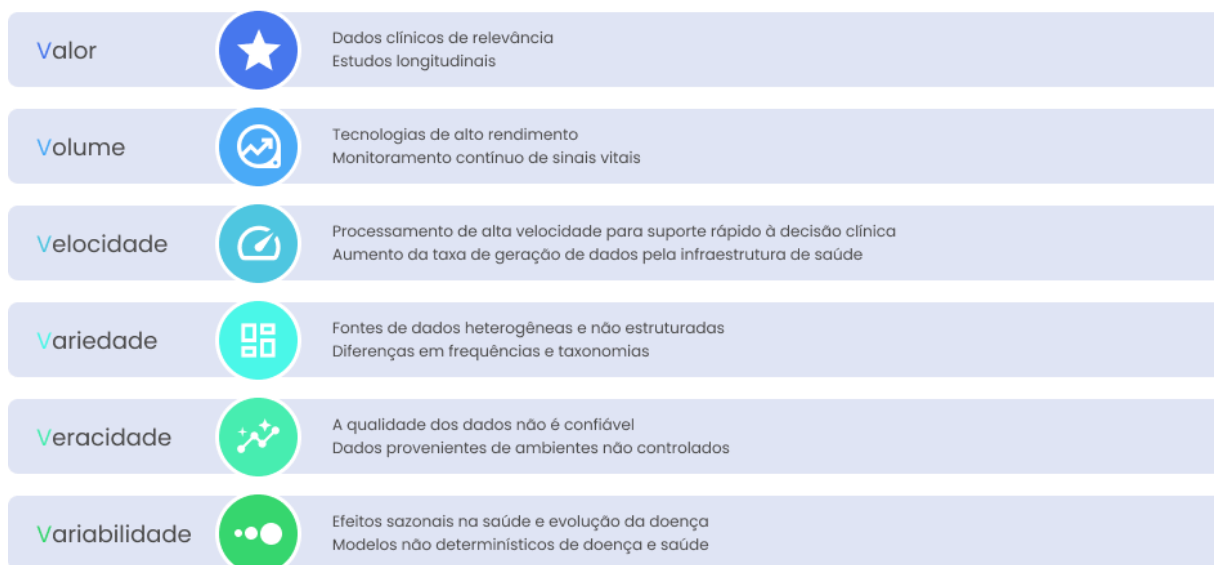
Na última década, a taxa de geração de dados cresceu de forma significativa em diferentes áreas da sociedade. Tal crescimento ocorreu devido ao avanço da tecnologia da informação (TI), de forma que o processo de gerar dados tornou-se muito mais fácil e constante (CHEN; MAO; LIU, 2014). Além disso, a evolução das tecnologias de computação em nuvem e internet das coisas segue favorecendo cenários no qual o volume de dados disponíveis é cada vez maior (CHEN; MAO; LIU, 2014). Nestes casos, a capacidade das arquiteturas e infraestruturas de TI tradicionais disponíveis em diversas empresas podem ser excedidas. Desta forma, cada vez mais o crescimento exponencial da quantidade de dados introduz novas problemáticas acerca de como armazenar e gerenciar conjuntos volumosos e heterogêneos de informação (CHEN; MAO; LIU, 2014).

A definição do termo *big data* é abstrata, pois outras características são levadas em consideração além dos altos volumes de dados encontrados nestes ambientes (CHEN; MAO; LIU, 2014). De forma geral, pode-se tratar *big data* como grandes conjuntos de dados que não podem ser adquiridos, gerenciados e processados por meio dos recursos tradicionais disponíveis de forma eficaz (CHEN; MAO; LIU, 2014). Além desta, existem ainda outras definições, como

a de Laney (2001) que define ambientes *big data* a partir de um modelo baseado em 3Vs, sendo estes: Volume, Velocidade e Veracidade. Neste modelo, Volume faz referência à geração de grandes quantidades de dados; Velocidade significa que estes dados devem ser processados o mais rápido possível, de forma a possibilitar maior extração de valor comercial; Variedade faz alusão aos diferentes tipos de dados encontrados em cenários *big data*, como dados estruturados, semiestruturados e não estruturados (CHEN; MAO; LIU, 2014). Empresas renomadas como IBM (ZIKOPOULOS; EATON, 2011) e departamentos da Microsoft (MEIJER, 2011) fizeram uso desta definição por tempo considerável. Com o passar dos anos, outros Vs foram adicionados à definição de Laney, como Valor, oriundo do relatório de Gantz (2011), no qual este enuncia que cenários *big data* são descritos como uma nova geração de tecnologias e arquiteturas, pensadas para extrair valor comercial de conjuntos de dados com grande volume e variedade, de forma a viabilizar capturas, descobertas e análises em alta velocidade. Posteriormente, Variabilidade e Veracidade foram incluídos ao conjunto de Vs, representando, respectivamente, a consistência dos dados obtidos ao longo do tempo e confiabilidade dos mesmos (ANDREU-PEREZ et al., 2015).

Um dos principais domínios em que contextos de *big data* tornaram-se comuns é o da saúde. Neste ambiente, as características apresentadas anteriormente para cenários *big data* tornam-se explícitas. Inicialmente, tem-se que o setor da saúde há muito tempo está habituado a lidar com grandes volumes de dados, normalmente partindo de zettabytes de informações (ANDREU-PEREZ et al., 2015). Em seguida, pode-se verificar a alta variedade dos dados presentes no meio, analisando as diferentes fontes pelas quais estes são gerados, como: testes laboratoriais, dados provenientes de sensores, matrizes de genes, textos livres, informações demográficas, entre outras. Desta forma, constitui-se um ambiente de dados heterogêneos, que se diferenciam em frequência e taxonomia (ANDREU-PEREZ et al., 2015). Além destes, tem-se ainda a característica de Velocidade, que se manifesta por meio da alta frequência de geração de dados por meio da infraestrutura de TI, presente em instituições de saúde, o que requer altos níveis de agilidade ao processar os dados gerados, a fim de auxiliar no processo de tomada de decisão clínica (ANDREU-PEREZ et al., 2015). Por fim, tem-se a principal característica de dados gerados nos domínios de saúde que é o alto valor agregado a estes, de forma que processá-los e analisá-los possibilita grandes melhorias na qualidade do serviço prestado ao paciente, além de maiores níveis de produtividade das equipes profissionais presentes em ambientes médicos (REDDY, 2015). Outros atributos de ambientes *big data*, como Veracidade e Variabilidade e seus espelhamentos no setor da saúde são exibidos na Figura 1.

**Figura 1 – Seis Vs de big data contextualizados para ambientes da saúde**



Fonte: Adaptado de (ANDREU-PEREZ et al., 2015)

## 2.4 Visual Data Mining

Pode-se definir *Visual Data Mining* como o processo de interação e descoberta de padrões por meio de representações visuais de determinado conjunto de dados (LIU; ÖZSU, 2009). Em outras palavras, este consiste em apresentar os dados a serem analisados de forma visual, de modo a possibilitar que o usuário final responsável pela análise tenha uma ampla visão dos dados, tire conclusões e interaja diretamente com o conjunto de informações analisadas. (KEIM, 2002). Atualmente, as técnicas de *Visual Data Mining* provaram-se recursos valiosos para o processo de análise exploratória de dados e, ainda, tem alto potencial para lidar com bases de dados altamente volumosas (KEIM, 2002). Neste contexto, realizar uma análise de dados por meio de técnicas visuais é particularmente útil em cenários onde o nível de conhecimento a respeito dos dados é baixo e os objetivos da análise ainda não foram bem definidos. Desta forma, caso necessário, realizar ajustes e mudanças nos objetivos da análise torna-se um procedimento automático, uma vez que o usuário está envolvido no processo (KEIM, 2002).

Para que o processo de *Data Mining* seja completamente efetivo, é de suma importância a inclusão do fator humano durante a exploração de dados, de forma a combinar habilidades humanas como flexibilidade, criatividade e conhecimentos gerais com a grande capacidade de



armazenamento e processamento dos computadores atuais (KEIM, 2002). Neste cenário, as técnicas de *Visual Data Mining* têm como objetivo realizar a inclusão do fator humano durante o processo de exploração dos dados, uma vez que estas empregam representações visuais das bases de dados (CHEN; GUO; WANG, 2015; KEIM, 2002), de forma a possibilitar que o responsável pela análise tenha percepções, conclusões e interaja diretamente com os dados analisados (KEIM, 2002). Além disso, o envolvimento direto do usuário final no processo de análise provê benefícios sobre as técnicas tradicionais de mineração de dados, como: técnicas de *Visual Data Mining* podem lidar com dados heterogêneos e que apresentam a presença de ruídos e técnicas de *Visual Data Mining* são mais intuitivas e não requerem conhecimentos prévios, como matemática ou algoritmos estatísticos (KEIM, 2002). Portanto, a utilização de *Visual Data Mining* permite explorar os dados de forma rápida e, ocasionalmente, fornece melhores resultados, especialmente em cenários em que os algoritmos tradicionais de *Data Mining* falham (KEIM, 2002).

Por fim, é importante mencionar que o processo de exploração visual de dados normalmente é composto por três etapas, que consistem em inicialmente obter uma visão geral dos dados, na qual o usuário é capaz de identificar os padrões de seu interesse (KEIM, 2002). Em seguida, para que o usuário possa analisar os padrões detectados, este deve ser capaz de realizar uma busca detalhada por meio de recursos de zoom e filtragem, a fim de obter uma visão mais detalhada dos dados, etapa também conhecida como detalhes sob demanda - do inglês, *Details On Demand* (KEIM, 2002). Tecnologias de visualização podem ser utilizadas durante todas as três etapas do processo de exploração (KEIM, 2002).

## **2.5 Classificação de técnicas de Visual Data Mining**

As técnicas de *Visual Data Mining* presentes na literatura podem ser classificadas de acordo com três critérios distintos: o tipo de dados a serem visualizados; a técnica de visualização empregada na análise; a forma de interação e distorção a ser utilizada durante o processo de análise (KEIM, 2002).

### **2.5.1 Dados a serem visualizados**

Na área de visualização, normalmente o volume de dados é elevado, de modo que cada registro possui suas especificidades, como número de variáveis e dimensionalidade (KEIM,

2002). O conceito de dimensionalidade dos dados está relacionado com o número de variáveis que este possui, de modo que tal característica pode ser classificada em unidimensional, bidimensional e multidimensional. Existem ainda formatos de dados mais complexos, como dados textuais. (KEIM, 2002).

Dados unidimensionais tendem a possuir apenas uma dimensão densa. Dados temporais são exemplos típicos dos dados incluídos nesta classe, uma vez que a cada ponto temporal, pode-se observar um ou mais valores associados (KEIM, 2002). Por sua vez, dados bidimensionais possuem duas dimensões distintas, e podem ser exemplificados por bases de dados geográficos, nas quais longitude e latitude são representadas por duas dimensões diferentes (KEIM, 2002). Tem-se ainda, dados multidimensionais. Estes são caracterizados por possuírem mais de três atributos e podem ser exemplificados pelas tabelas pertencentes a bases de dados relacionais, as quais normalmente possuem dezenas ou centenas de atributos (KEIM, 2002). Dados pertencentes a esta classificação são mais complexos de serem visualizados, pois não aceitam visualizações tradicionais como os dados com até três dimensões (KEIM, 2002).

Por fim, existem dados que não podem ser classificados em termos de quantidade de dimensões, como dados textuais. Dados que não podem ser descritos em termos do número de dimensões são difíceis de serem visualizados, pois estes não podem ser representados por meio de números. Desta forma, a maior parte das técnicas de visualização não podem ser aplicadas a dados textuais (KEIM, 2002). Nestes casos, processos de transformação de dados são utilizados, a fim de possibilitar descrevê-los por meio de números (KEIM, 2002).

### **2.5.2 Técnicas de visualização**

Existem diversas técnicas de visualização disponíveis na literatura. Além das tradicionais técnicas destinadas a dados unidimensionais e bidimensionais, como gráficos de eixo X e Y (X, Y e Z para dados tridimensionais), gráficos de barra e gráficos de linha, existem ainda técnicas de visualização mais aprimoradas (KEIM, 2002). Telas transformadas geometricamente – do inglês, *Geometrically Transformed Displays* – tem como objetivo apresentar bases de dados multidimensionais de forma visual para encontrar transformações pertinentes à análise (KEIM, 2002). Esta classe de técnicas é constituída de métodos de estatística exploratória, como gráficos de dispersão – do inglês, *Scatterplots* (KEIM, 2002). Existem ainda, técnicas de visualização empregadas em bases de dados multidimensionais que visam representar o valor de determinado registro como atributos de ícones, que podem ser rostos, estrelas, bonecos de palito, entre outros. Estas são conhecidas como telas icônicas – do

inglês, *Iconic Displays* (KEIM, 2002). Por sua vez, a técnica de tela de pixel denso – do inglês, *Dense Pixel Displays* – visa mapear o valor de determinada dimensão para um *pixel* colorido. Posteriormente, estes serão agrupados em pixels pertencentes a cada dimensão em áreas adjacentes. Esta técnica é particularmente útil para visualizar conjunto de dados volumosos, uma vez que se utiliza um pixel para cada valor (KEIM, 2002). Por fim, pode-se citar as técnicas de exibição empilhada – do inglês *Stacked Displays*, que visam apresentar determinado conjunto de dados de forma particionada e hierárquica (KEIM, 2002).

### 2.5.3 Técnicas de interação e distorção

Para que o processo de exploração de dados possa ser efetivo, o analista deve ser capaz de interagir com os dados a ele apresentados (KEIM, 2002). Para isso, existem técnicas de interação, que possibilitam ao analista de dados interagir diretamente com a visualização e alterá-la dinamicamente, de acordo com os objetivos do processo de exploração (KEIM, 2002). Além disso, técnicas de visualização permitem a combinação de diferentes visualizações. (KEIM, 2002).

Por sua vez, técnicas de distorção fazem com que o analista de dados seja capaz de manter-se focado em determinado subconjunto de dados, enquanto a visão geral do conjunto de dados original é mantida (KEIM, 2002). Para tal, utiliza-se recursos de filtragem interativa, nos quais por meio da interface gráfica permite-se que determinada porção dos dados seja diretamente selecionada para análise detalhada, ou ainda, pode-se selecionar os dados desejados através de consultas – do inglês, *queries* (KEIM, 2002).

## 2.6 Computação em nuvem

Computação em nuvem – do inglês, *Cloud Computing* pode ser definido como o processo de prover recursos de computação sob demanda por meio de internet. Portanto, não pode ser considerada como uma nova tecnologia, mas sim, uma nova maneira de oferecer recursos computacionais (KUO, 2011). *Cloud Computing* é, por muitos, considerada como a próxima etapa da evolução da computação distribuída, uma vez que tal modelo tem como objetivo utilizar recursos distribuídos de forma eficiente, de modo que seja possível alcançar altas taxas de transferência, necessárias para que infraestruturas capazes de lidar com problemas de escalabilidade sejam fornecidas via internet (RIMAL; CHOI; LUMB, 2009).

A maior adesão de *Cloud Computing* pode oferecer diversos benefícios para inúmeras áreas de operação da sociedade, uma vez que todas necessitam de evolução contínua e

sistemática, ao passo que os custos devem se manter otimizados (RIMAL; CHOI; LUMB, 2009). Deste modo, setores como saúde e cuidados médicos podem ter a qualidade de seus serviços aperfeiçoados por meio do emprego de *Cloud Computing*, como estimam especialistas da área (RIMAL; CHOI; LUMB, 2009). Vantagens como redução de gastos iniciais para a implantação de sistemas gerenciadores de Prontuário Eletrônico do Paciente e maior disponibilidade de recursos computacionais capazes de lidar com os complexos problemas de processamento de dados da área podem ser apontados como incentivos válidos para o emprego de computação em nuvem em contextos médicos (RIMAL; CHOI; LUMB, 2009).

Existem diferentes recursos fornecidos por serviços de *Cloud Computing* que podem ser de suma importância para a área médica (WORKNEH; ADEM; PRADHAN, 2018). Entre estes, pode-se citar a maior tolerância a falhas. Plataformas de *Cloud Computing* são capazes de fornecer aplicações com altos níveis de disponibilidade, pois uma vez que os recursos utilizados para servi-las encontram-se distribuídos pela internet. Em caso de falha, existirão outras instâncias da aplicação prontas para assumir o lugar da principal, sem qualquer tipo de interrupção (RIMAL; CHOI; LUMB, 2009). Além disso, outro recurso importante em âmbitos médicos é a segurança dos dados (WORKNEH; ADEM; PRADHAN, 2018), recurso este, amplamente desenvolvido em plataformas de *Cloud Computing*, uma vez que a segurança dos dados armazenados na plataforma, além de uma questão técnica, é, por vezes, questão contratual. Desta forma, constitui-se um cenário no qual aplicações pertencentes a domínios médicos podem armazenar, teoricamente, quantidades ilimitadas de dados em uma plataforma distribuída que garante a segurança, alta disponibilidade, poder de processamento e tolerância a falhas.

## 2.7 Microsserviços

Microsserviços podem ser definidos como módulos coesos de determinada aplicação e independentes, capazes de se comunicar por meio de mensagens (DRAGONI et al, 2017). Por consequência, a arquitetura de microsserviços é definida como uma aplicação distribuída onde todos os módulos por esta utilizados, encontram-se no formato de microsserviço (DRAGONI et al, 2017). Com isso, aplicações construídas com base nesta arquitetura tendem a otimizar benefícios disponibilizados por plataformas de *Cloud Computing*, como alta disponibilidade, pois uma vez que todos os módulos da aplicação são distribuídos, é possível reduzir o tempo de implantação de novas versões do sistema, de forma a otimizar a disponibilidade da aplicação como um todo (LARRUCEA et al, 2018). Além disso, aplicações baseadas em microsserviços

oferecem ainda, maior reuso de *software*, fator de relevância como mostram pesquisas realizadas no campo da Engenharia de Software (DA SILVA et al, 2019), pois, uma vez que determinada plataforma possui componentes reutilizáveis, esta tende a possuir um processo de manutenção menos complexo e custoso (KHOSHKBARFOROUSHHA et al, 2014).

Fundações como a OpenEHR – do inglês, *Open Eletronic Health Record*, cada vez mais incentivam o emprego de técnicas que visam o reuso de software, como microsserviços, durante a comunicação entre aplicações da área da saúde (DA SILVA et al, 2019). Em contrapartida, pesquisadores mostram que em grande parte das aplicações disponíveis no domínio da saúde, são utilizadas arquiteturas com alto nível de acoplamento que não disponibilizam seus componentes de forma independente, conhecidas como monolitos (DA SILVA et al, 2019). Deste modo, os benefícios do uso arquiteturas baseadas em microsserviços em conjunto com plataformas de *Cloud Computing* podem ser melhor aproveitados por aplicações inseridas em contextos médicos.

## 2.8 Trabalhos correlatos

São inúmeros os trabalhos na área de *Visual Data Mining*. Para ilustrar a sua relevância e atualidade, alguns dos estudos que podem ser citados são os realizados por Shaban-Nejad et al. (2017), Ruan et al. (2018), Tang et al. (2019) e Barroca Filho et al. (2020). Os principais critérios de escolha para os trabalhos a seguir apresentados foram:

- Revistas de publicação: foram selecionados apenas trabalhos de revistas qualificadas e de alto fator de impacto.
- Ambientes de visualização: foram considerados apenas trabalhos que lidam diretamente com a incorporação de técnicas de visualização de dados em plataformas voltadas para ambientes da saúde.
- Ano de publicação dos trabalhos: foram considerados apenas trabalhos relativos aos últimos 4 anos, de modo que fosse possível observar a evolução do tema na literatura científica.

### 2.8.1 Ambiente de visualização de dados heterogêneos da área da saúde

Shaban-Nejad et al. (2017) propõem em seu trabalho o Registro de Saúde da População — PopHR, do inglês Population Health Record — uma plataforma que, por meio da integração de dados heterogêneos, cálculos de indicadores de saúde e interpretação de indicadores para uma determinada população, visa facilitar as tomadas de decisão que permeiam a saúde da

população. Sua proposição tem o objetivo de auxiliar os profissionais da área médica, bem como os responsáveis por formular políticas para a área da saúde, a monitorar a saúde da população e o funcionamento do sistema de saúde, além de possibilitar pesquisas nessa área. Isto é feito através da integração e exploração destes dados médicos, provenientes de diversas fontes como dados administrativos, registros clínicos, entre outras. Tanto a integração quanto a exploração são processos realizados de forma automática e sobre um elevado volume de dados.

De modo a entender melhor o PopHR, é importante mencionar a teoria da Saúde Pública Baseada em Evidências — EBPH, do inglês Evidence-Based Public Health — (KOHATSU; ROBINSON; TORNER, 2004). EBPH pode ser entendida como a junção de pesquisas científicas, representadas por intervenções, com as preferências da população para melhorar a saúde da mesma. Tem-se que o alinhamento e associação das informações sobre uma população baseadas em evidências relevantes, aprimora o processo de tomada de decisões realizado por esta população. Dentre os tipos de evidência identificados, tem-se que o PopHR utiliza-se do tipo 1: conhecimento sobre as relações causais entre as variáveis. Este é utilizado de modo que a plataforma identifica conceitos relevantes na saúde da população, como condições de vida e saúde, a fim de gerar indicadores que auxiliem os usuários do sistema de saúde a identificarem fatores relevantes para a saúde da população, bem como para o próprio sistema.

A projeção da arquitetura do PopHR foi pensada de modo que esta pudesse exibir indicadores em nível de população, fossem esses pertinentes ou não à área da saúde, além de realizar automaticamente a geração de indicadores utilizando definições de cases e dados heterogêneos, além de integrá-los, utilizando o conhecimento do domínio, e, por fim, a visualização e comparação desses indicadores em parcelas da população. Além disso, de modo a garantir a confidencialidade, responsabilidade e transparência do sistema, utilizou-se uma arquitetura cliente-servidor, somada a um *back-end* para processamento de dados e um *front-end* de visualização.

É importante ressaltar que o PopHR possui suporte para realização da coleta, normalização, alinhamento, integração e transformação de dados estruturados e não estruturados, advindos de diversas fontes.

### **2.8.2 Ambiente de visualização de condições médicas extraídas de notas clínicas**

Em seu trabalho, Ruan et al. (2018) propõem uma forma de lidar com os desafios enfrentados na interpretação de dados contidos em notas clínicas — componente importante

dos PEP. Tais desafios consistem, basicamente, nas dificuldades para determinar condições médicas de um paciente ou encontrar informações de forma rápida e eficaz, devido ao elevado volume de dados, em sua maioria, textuais. De forma a tentar resolver esses problemas, são propostos uma interface gráfica projetada para permitir a visualização de um resumo das informações médicas concernentes aos pacientes, bem como um sistema que viabiliza a extração e visualização de informações a partir de notas clínicas estruturadas.

Os registros médicos, que possuem as mais variadas classes, podem ser visualizados a partir de duas interfaces: uma baseada em registros espaciais, isto é, em qual região corporal concentra-se a enfermidade do paciente, e outra para registros temporais. Para a primeira, são definidos seis sistemas fisiológicos, de modo que a partir da extração das informações sobre a doença e a região corporal em que ela se encontra, os dados são representados espacialmente, indicados através de uma área circulada na representação visual do sistema fisiológico em que a doença está situada. Já a representação temporal, baseia-se na elaboração de uma linha temporal, contendo informações como medicamentos ministrados, testes laboratoriais e tratamentos, de forma a mostrar a evolução clínica do paciente.

É importante destacar que as notas clínicas passam por um pré-processamento, o qual consiste na extração do texto por meio da utilização da ferramenta Metamap (METAMAP, 2021), capaz de reconhecer vocábulos presentes no Sistema de Linguagem Médica Unificado — UMLS, do inglês Unified Medical Language System — (NLM, 2021), seguida da conversão do texto plano em formato XML (eXtensible Markup Language), por fim, utiliza-se o Apache OpenNLP (APACHE OPENNLP, 2021) — biblioteca utilizada para processamento de linguagem natural — para treinar o modelo de entidade de detecção, de modo que esta possa identificar e categorizar registros temporais.

### **2.8.3 Ambiente de visualização do progresso da doença de obstrução pulmonar crônica**

No trabalho realizado por Tang et al. (2019), há a proposição de uma abordagem contendo duas etapas, que visa permitir a visualização de dados pertinentes à progressão de doenças pulmonares, mais especificamente a doença de obstrução pulmonar crônica — COPD, do inglês, *Chronic Obstructive Pulmonary Disease* — a partir de notas clínicas não estruturadas.

As etapas para o desenvolvimento do projeto foram estabelecidas da seguinte forma: desenvolvimento de um modelo de *deep learning* contendo quatro camadas, de modo a estender

a arquitetura de *Long-short Term Memory* — LSTM, traduzido livremente como “memória de longo-curto prazo” — a fim de configurar de forma apropriada os intervalos de tempo e possibilitar a representação de dados de séries temporais com amostragem irregular. E, a criação de uma visualização temporal, alicerçada nos intervalos de tempo definidos anteriormente, de modo a exibir a progressão da COPD.

O resultado da primeira etapa, isto é, o modelo de *deep learning*, tem, conforme mencionado anteriormente, quatro camadas. A primeira é destinada ao pré-processamento e incorporação de palavras a fim de preparar os dados. Nela, é realizado um processo de codificação para cada amostra de dados, seguido de um balanceamento dos dados, efetuado por meio de um preenchimento das amostras, de modo a garantir que elas possuam o mesmo comprimento. Por último, é realizada a incorporação de palavras, de modo que as palavras extraídas são codificadas a fim de serem representadas por um vetor denso em um espaço vetorial contínuo. A segunda etapa baseia-se na LSTM e é utilizada para prever a data de morte do paciente, uma vez que essa arquitetura é adequada para predições, posto que suporta os intervalos de tempo desiguais entre eventos. As terceira e quarta camadas, *flatter* e *dense* — traduzidas livremente como “plana” e “densa”, respectivamente —, são as responsáveis por apreender as janelas de tempo que possuem amostragem irregular.

Na etapa de visualização, é utilizada uma linha do tempo em espiral, uma vez que essa se mostra ideal para a representar a progressão da doença. Os fatores que corroboram com a decisão de utilizá-la são o fato de esta realizar a exposição de forma compacta de períodos longos de tempo e evitar que a correlação entre dois eventos com parâmetros similares seja perdida. Utiliza-se também um mapa geográfico agregado à linha do tempo em espiral, uma vez que, conforme explicado no trabalho de Hewagamage, Hirakawa e Ichikawa (1999), essa junção possibilita representar padrões temporais de eventos considerando atributos espaciais.

#### **2.8.4 Ambiente de visualização de dados do paciente em forma de dashboard**

Por fim, tem-se o trabalho de Barroca Filho et al. (2020), no qual os autores apresentam as técnicas utilizadas para a elaboração de uma *dashboard* para a plataforma SigSaúde, que é um sistema de prontuário eletrônico de saúde que contém as informações dos pacientes que utilizam os serviços de saúde oferecidos pela Universidade Federal do Rio Grande do Norte. Uma vez que compreende informações como o tipo sanguíneo, estado civil e etnia de cada paciente, tal qual um mapa que possibilita verificar o número de usuários listados por cidade, a



*dashboard* visa possibilitar, por meio de gráficos interativos, a visualização de informações de pacientes, de modo a permitir que análises sejam realizadas. Por esse motivo, favorece o processo de tomada de decisão, uma vez que auxilia na interpretação dos dados pelos pesquisadores e demais responsáveis da área.

Salienta-se que para um melhor entendimento do trabalho proposto por Barroca Filho et al., é necessário entender os softwares que estes empregaram para a sua execução. Para o desenvolvimento da *dashboard*, os autores utilizaram a *ELK Stack* (traduzido livremente como Pilha ELK), na qual as três letras são as iniciais das ferramentas usadas para construí-lo, são elas: *Logstash* (ELASTIC, 2021a), que promove a união de múltiplas fontes de dados, a exemplo de bancos de dados e *log files* — registros de eventos de um dado sistema —, de modo a unificá-los e, assim, padronizar os dados; *ElasticSearch* (ELASTIC, 2021b), cuja funcionalidade principal é a indexação de dados, de modo a possibilitar consultas de informações de forma eficaz e escalonável, por meio de requisições do Protocolo de Transferência de Hipertexto — HTTP, do inglês, *Hypertext Transfer Protocol* — através de Interface de Programação de Aplicações — API, do inglês, *Application Programming Interface* — alicerçada em Transferência de Estado Representacional — REST, do inglês, *Representational State Transfer*; por último, tem-se a Kibana (ELASTIC, 2021c), utilizada para a visualização dos dados, uma vez que oferece recursos para a criação de gráficos, o que facilita a interpretação dos dados por parte dos usuários, seja na associação de informações ou na identificação de padrões.

Assim, o desenvolvimento da *dashboard* foi separado em etapas, que podem ser norteadas a partir das ferramentas utilizadas. Para as etapas de extração dos dados da plataforma SigSaúde, obtenção da localização geográfica — para posterior mapeamento de pacientes por cidade — e anonimização dos dados, utilizou-se o *Logstash*. Após, para a indexação dos dados e geração de uma interface para a realização de consultas, recorreu-se à *Elasticsearch*. Por fim, para a construção das visualizações e do *dashboard* em si, empregou-se a Kibana.

## 2.9 Principais características dos trabalhos correlatos

Por meio da análise dos trabalhos correlatos descritos, foi realizada uma comparação entre as principais características da literatura acadêmica apresentada. Esta, é apresentada no Quadro 1. Uma vez analisada, torna-se perceptível a existência de uma lacuna na literatura científica, de modo que não existem ambientes de visualização capazes de lidar com as diferentes características dos dados presentes no setor da saúde ao passo em que possuem

arquiteturas que utilizam de recursos que otimizam os fatores de escalabilidade, reduzem os custos da operação e ainda, permitem maiores níveis de reuso de software para o ambiente desenvolvido.

**Quadro 1 - Comparativo dos trabalhos correlatos**

	(SHABAN-NEJAD et al., 2017)	(RUAN et al., 2018)	(TANG et al., 2019)	(BARROCA FILHO et al., 2020)
<b>Capaz de visualizar dados estruturados</b>	✓	✗	✗	✓
<b>Capaz de visualizar dados não estruturados</b>	✓	✓	✓	✗
<b>Utiliza recursos de <i>Cloud Computing</i></b>	✗	✗	✗	✓
<b>Utiliza arquitetura baseada em microserviços</b>	✗	✗	✗	✗

Fonte: Elaborado pelo autor.

## 2.10 Considerações finais

Neste capítulo, foram apresentados os principais conceitos PEP e como este tornou-se recurso importante para o setor da saúde. Além disso, discutiu-se sobre *Big Data* e suas características. Em seguida, conceitos de Visual Data Mining e suas classificações foram abordados. Por fim, discorreu-se sobre *Cloud Computing* e como esta passa a impactar áreas como a saúde.

Por fim, foram abordados quatro trabalhos correlatos com o intuito de apresentar o estado atual da arte e ilustrar quais foram as contribuições sobre o tema nos últimos 4 anos e quais lacunas podem ser preenchidas.

# Capítulo 3

## Desenvolvimento

### 3.1 Considerações iniciais

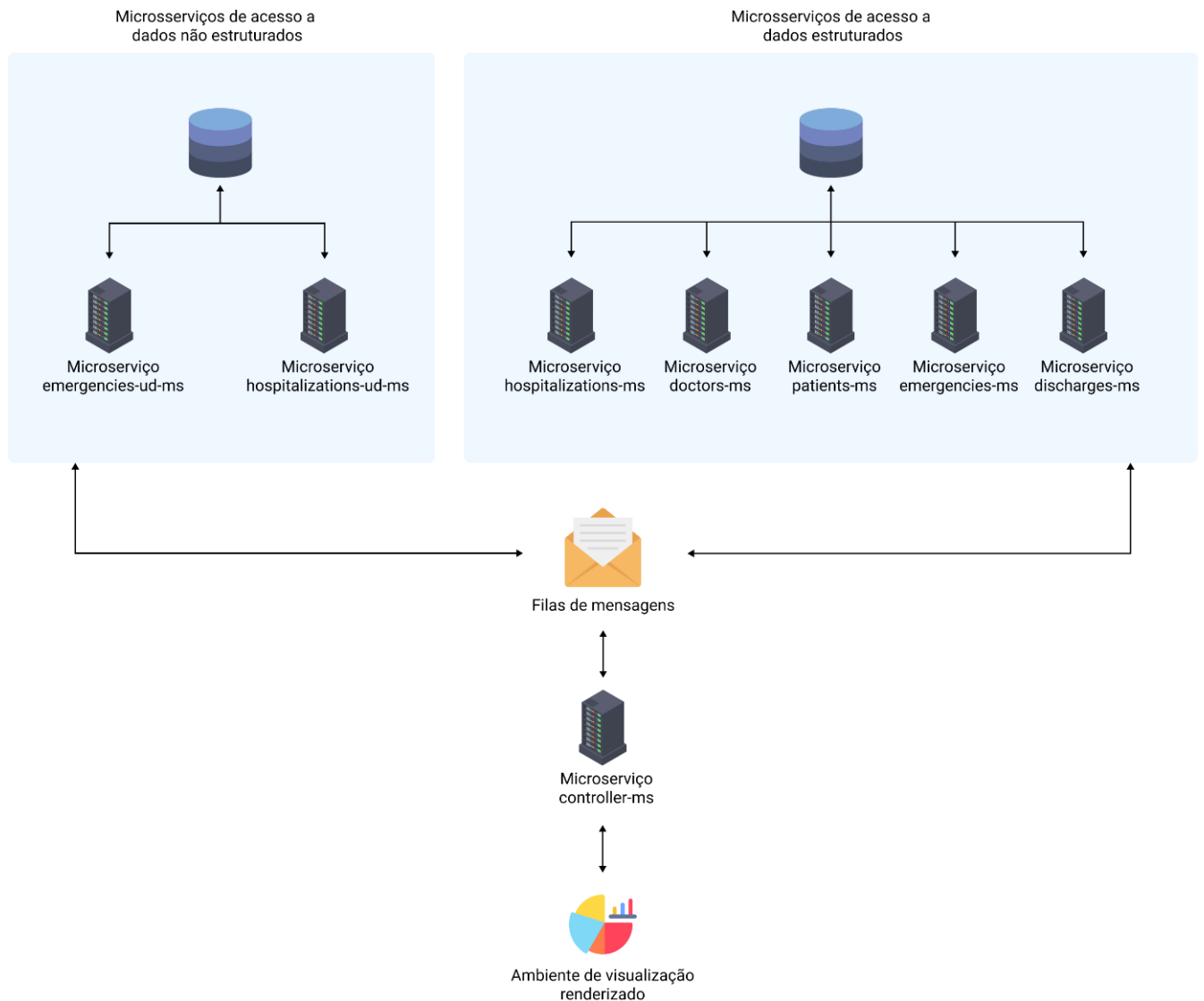
Neste capítulo, é apresentado o processo de desenvolvimento do ambiente de visualização de dados estruturados e não estruturados. Além disso, a arquitetura de microsserviços em que a aplicação se baseia está detalhada, de modo a esclarecer aspectos técnicos de implementação, como linguagens e plataformas, além de elucidar como cada componente foi integrado com os demais.

### 3.2 Arquitetura do sistema

Na Figura 2 está ilustrada a arquitetura baseada em microsserviços hospedados em *cloud*, desenvolvida para o visualizador. Como evidencia-se, o ponto inicial de todas as chamadas feitas para o sistema a partir do ambiente de visualização será o serviço controlador. Este possui responsabilidades, como autenticar o responsável pela chamada, encaminhá-la para o serviço correto e devolver a resposta esperada. Para isso, o controlador necessita de uma base de dados, que armazenará as informações dos usuários do sistema.

Além do serviço controlador, existem ainda os serviços responsáveis por fornecer os dados de internações, emergências, pacientes e consultas do hospital. Estes são diretamente conectados a base de dados do hospital e são responsáveis por gerenciar individualmente cada uma das entidades citadas anteriormente. Por fim, vale ressaltar que a comunicação entre os microsserviços citados foi implementada utilizando serviços de mensageria.

**Figura 2 – Arquitetura do sistema**



Fonte – Elaborado pelo autor.

### 3.3 Microserviço controlador

O serviço controlador é o principal componente do sistema, de forma que sua função fundamental é disponibilizar todos os recursos ofertados pelos microserviços do hospital por meio de uma interface REST. Desta forma, todos os recursos da aplicação tornam-se acessíveis por meio de um único contrato de comunicação. No Quadro 2 estão contidos a interface disponibilizada pelo controlador e seu respectivo mapeamento.

**Quadro 2 – Interfaces disponibilizadas pela aplicação**

Método HTTP	Interface de acesso	Microserviço
POST	/users	controller-ms
POST	/authentication	controller-ms
GET	/patients	patients-ms
GET	/hospitalizations	hospitalizations-ms
GET	/hospitalizations/:id	hospitalizations-ud-ms
GET	/emergencies/	emergencies-ms
GET	/emergencies/:id	emergencies-ud-ms
GET	/discharges	discharges-ms
GET	/doctors	doctors-ms

Fonte: Elaborado pelo autor

Além disso, também foram implementados no controlador funcionalidades importantes, como autenticação para controle de acesso, recurso indispensável para sistemas de saúde, e a integração com serviços de mensageria para a comunicação entre os microserviços.

### 3.3.1 Autenticação

A autenticação da aplicação foi construída baseada no protocolo JWT — do inglês, *JSON Web Token* —, forma compacta e segura de representar e transmitir informações entre duas partes, como permissões e direitos de acesso (JONES; BRADLEY; SAKIMURA, 2015). Desta forma, é possível transitar informações a respeito do usuário autenticado e seus direitos de acesso por meio de uma estrutura JSON — *JavaScript Object Notation*, traduzido livremente como “notação de objeto JavaScript” — confiável que pode ser verificada, uma vez que esta foi digitalmente assinada. Para tal, foi utilizado o algoritmo HMAC — *Hash-based Message Authentication Code*, traduzido livremente como “código de autenticação de mensagens baseado em *hash*” — em conjunto com uma base de criptografia secreta. Na Figura 3 está ilustrada a estrutura de um JWT.



### 3.4.1 Microserviços de dados estruturados

Constituem os microserviços de acesso a dados estruturados: *patients-ms*, *hospitalizations-ms*, *emergencies-ms*, *discharges-ms* e *doctors-ms*. Ambos foram implementados por meio da plataforma .NET 5 com a linguagem orientada a objeto C#.

Além disso, é importante citar que o banco de dados que contém as informações estruturadas a respeito de pacientes, internações, atendimentos de emergência e médicos é o mesmo. Trata-se de um banco de dados relacional PostgreSQL.

Cada um dos componentes citados anteriormente trata-se de uma aplicação integrada à fila de mensagens que, inicialmente, assume um estado de *consumer*, isto é, a mesma passa a monitorar constantemente a fila de comunicação à espera da chegada de uma nova mensagem. Neste momento, a mensagem é processada pela aplicação, que busca os recursos solicitados. Ao término do processamento da mensagem, os serviços passam a atuar como *producers*, de modo a enviar uma mensagem de *response* para a fila de comunicação do microserviço controlador, que, por sua vez, a devolverá para o componente responsável por realizar a chamada do recurso.

### 3.4.2 Microserviços de dados não estruturados

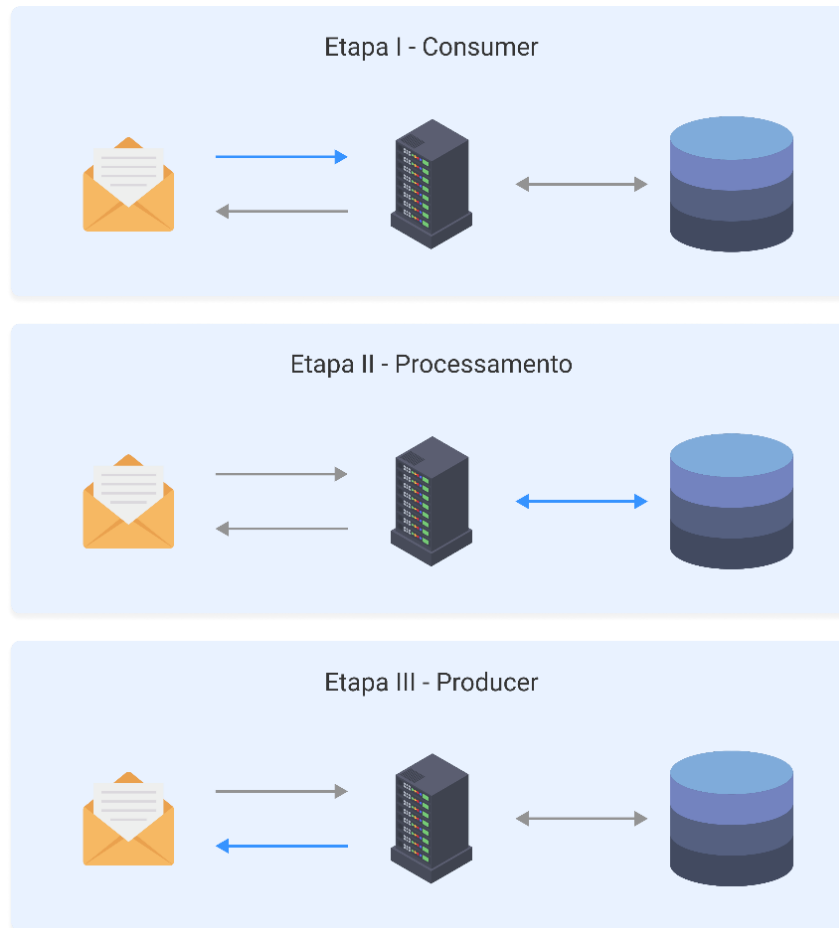
Por sua vez, tratam-se de microserviços de acesso a dados não estruturados: *hospitalizations-ud-ms* e *emergencies-ud-ms*. Estes componentes são responsáveis por acessar informações provenientes das notas clínicas de internações e atendimentos de emergência, respectivamente. Para isso, utiliza-se uma ontologia baseada no CID10 — Classificação Internacional de Doenças, décima versão —, na qual tais dados são processados e posteriormente, integrados para um armazém de dados relacional PostgreSQL.

Além disso, é importante citar que para a implementação do código fonte dos microserviços de acesso a dados não estruturados foi utilizada a linguagem de programação Java em sua versão 8.

Analogamente aos serviços de acesso a dados estruturados, foi construída uma aplicação que em primeiro momento, assume o estado de *consumer*, de forma a manter-se em estado de monitoramento da fila de mensagens e, após a chegada de uma nova mensagem, o processamento da mesma é iniciado. Outra similaridade baseia-se no fato de que ao terminar o processamento e aquisição dos dados, o componente muda sua atuação de *consumer* para

*producer*, de forma a produzir e enviar uma mensagem de *response* para a fila de comunicação do microserviço controlador. Este fluxo de execução que é comum para microserviços de acesso a dados estruturados e não estruturados está ilustrado na Figura 4 a seguir:

**Figura 4 – Fluxo de execução dos microserviços**



Fonte – Elaborado pelo autor.

### 3.5 Ambiente de visualização

A presente seção tem como objetivo apresentar o ambiente de visualização após implementado. O ambiente é responsável por consumir os dados disponibilizados pelos microserviços anteriormente apresentados e renderizá-los de forma gráfica.

Por se tratar de uma aplicação *web*, em sua implementação, foi utilizada a tecnologia SPA — *Single Page Application*, traduzido livremente como “aplicação de página única”. Esta consiste em uma página individual que pode ser atualizada de forma independente à ação de determinado usuário (JOSEPH, 2015). Desta forma, não é necessário que a página inteira seja recarregada como em aplicações *web* tradicionais. Isto ocorre pois todo o conteúdo necessário

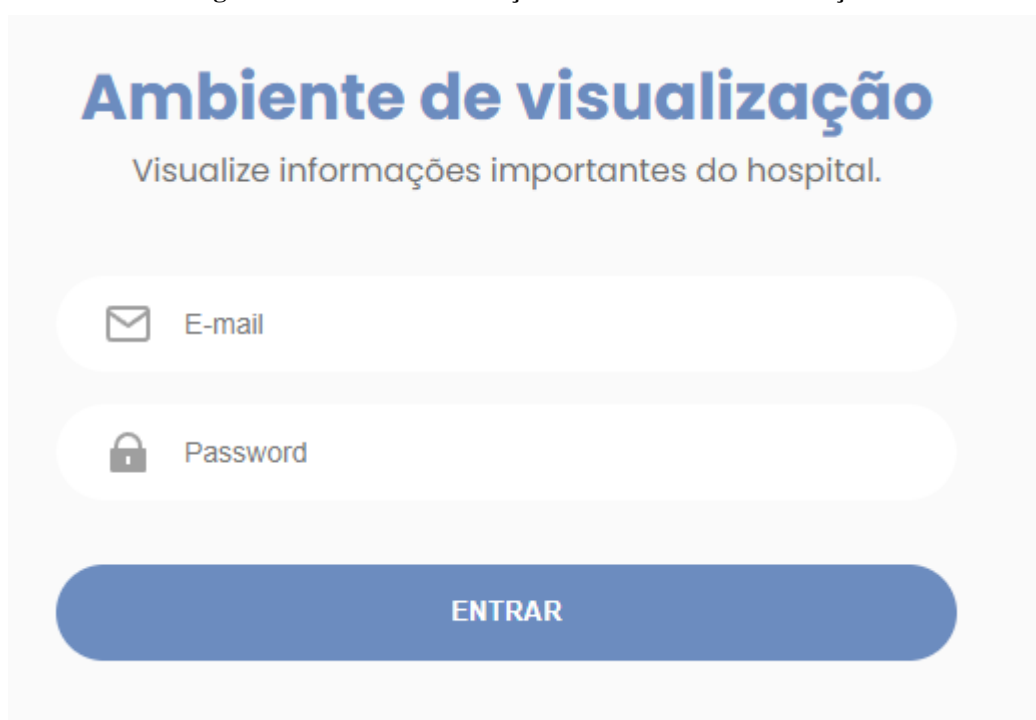


para a plena execução da aplicação é carregado no *browser* durante o primeiro acesso à página. Com isso, os níveis de interatividade, capacidade de resposta e satisfação do usuário são significativamente aprimorados (MESBAH; VAN DEURSEN, 2007).

### 3.5.1 Página de autenticação

Na Figura 5 está ilustrada a primeira página exibida ao usuário, no momento em que a aplicação é carregada. Nesta, o usuário deve autenticar-se com seus dados de acesso ao sistema para, então, ser direcionado para a página inicial. Caso o usuário não possua cadastro prévio na plataforma, seu acesso é negado e uma mensagem de erro é exibida em tela. Uma vez que o usuário se autentique com suas credenciais corretas, o mesmo é direcionado para a página inicial do ambiente de visualização

Figura 5 – Tela de autenticação do ambiente de visualização



A imagem mostra a tela de autenticação de um sistema. No topo, o título "Ambiente de visualização" está em azul escuro, com o subtítulo "Visualize informações importantes do hospital." em cinza. Abaixo, há dois campos de entrada: o primeiro, rotulado "E-mail" com um ícone de envelope, e o segundo, rotulado "Password" com um ícone de cadeado. Ambos os campos são brancos com bordas arredondadas. Na base da tela, há um botão azul escuro com o texto "ENTRAR" em branco.

Fonte – Elaborado pelo autor.

### 3.5.2 Página Home

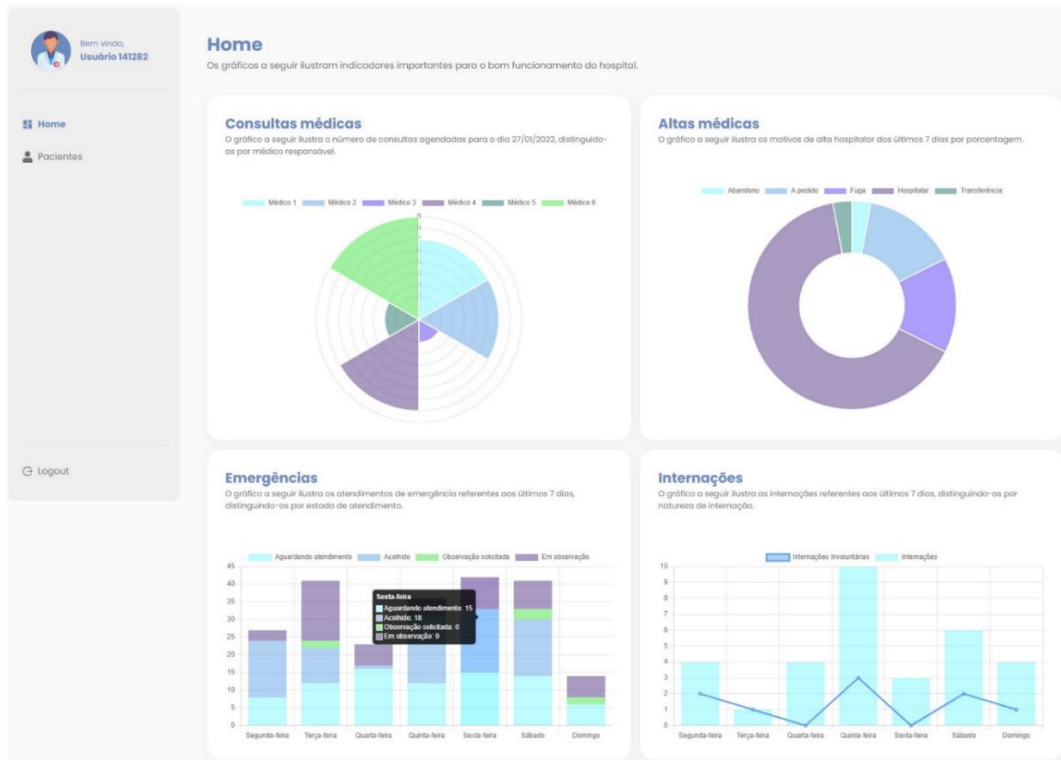
Ilustrada na Figura 6, a página inicial — ou *home* —, tem como intuito apresentar informações gerais do hospital relativas a curtos períodos de tempo, como: o gráfico de consultas médicas, que exhibe o número de consultas para cada um dos médicos do hospital no

dia atual; o gráfico de altas médicas, que informa o percentual das distintas justificativas para altas hospitalares dos últimos sete dias; o gráfico de emergências, responsável por informar o número de atendimentos de emergência realizados nos últimos sete dias, distinguindo-os por seus diferentes estados de atendimento; e o gráfico de internações, o qual ilustra o número de internações relativo aos últimos sete dias, onde destas, destacam-se ainda quais se tratam de internações involuntárias.

Além disso, é importante salientar que todos os gráficos presentes no ambiente de visualização possuem métodos para manipulação. Ao passar o mouse sobre qualquer uma das frações do gráfico, a cor desta se sobressai em relação às demais, destacando-a para o usuário. Além disso, ao clicar em um dos rótulos presentes na legenda do gráfico, sua fração correspondente é ocultada da visualização. Como abordado no capítulo 2, tais recursos de manipulação são de fundamental importância para os usuários da aplicação. Uma amostra destes recursos está ilustrada na Figura 11.

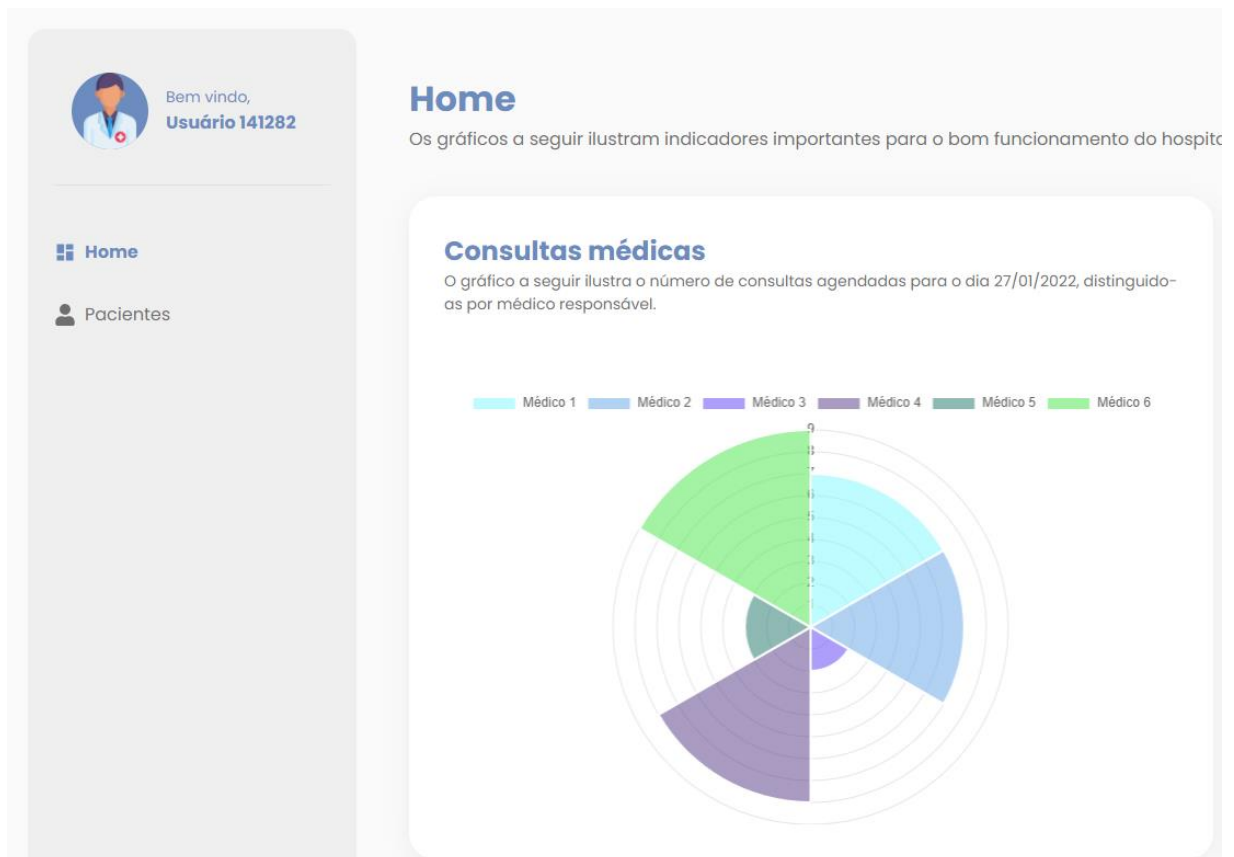
Vale ressaltar que toda a estrutura do ambiente de visualização é a mesma para as demais páginas da aplicação, ou seja, consiste em uma barra lateral à esquerda das páginas, que contém opções de acesso à página *home* e à página de pacientes. Além disso, a barra lateral oferece ao usuário ainda a opção de realizar *logout* na aplicação, de modo a retornar para a página de autenticação. Note que a barra lateral preenche todo o espaço vertical de visualização do usuário, porém, nas figuras a seguir, a mesma encontra-se suspensa no canto superior esquerdo da página. Isto ocorre, pois, todo o conteúdo visível, presente nas imagens, que se encontra abaixo da barra lateral, é visualizável na aplicação apenas por meio de rolagem de página.

Figura 6 – Tela inicial do ambiente de visualização



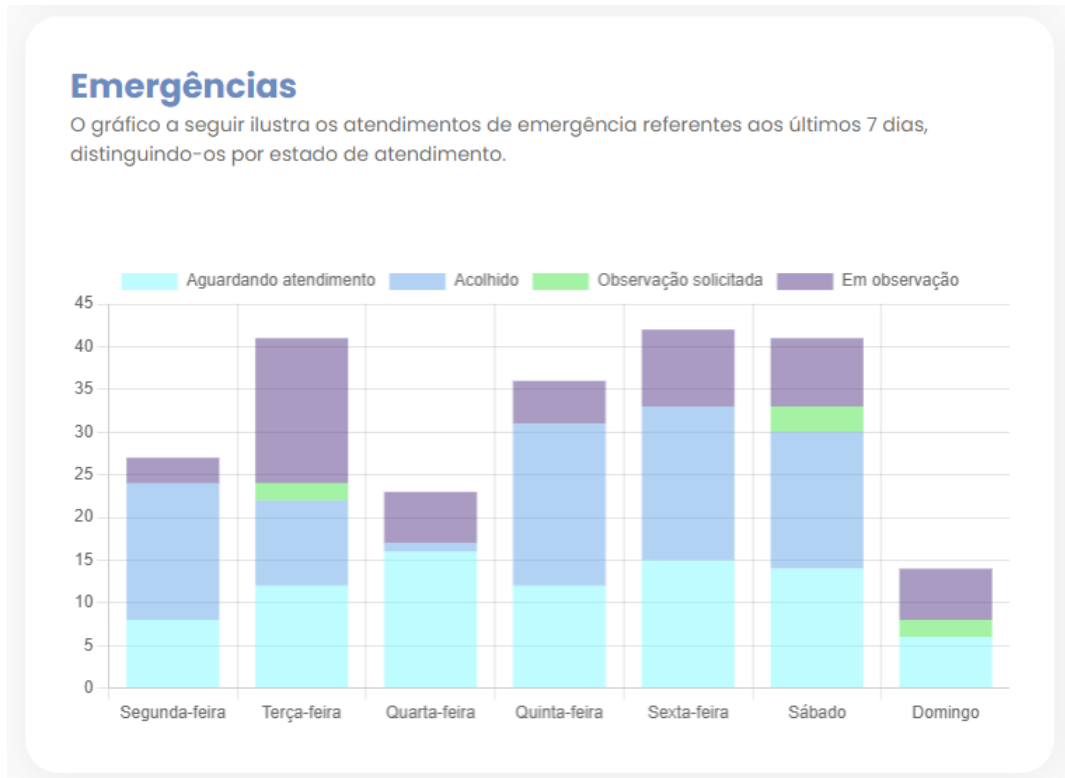
Fonte: Elaborado pelo autor.

Figura 7 – Gráfico de consultas médicas presente na página Home



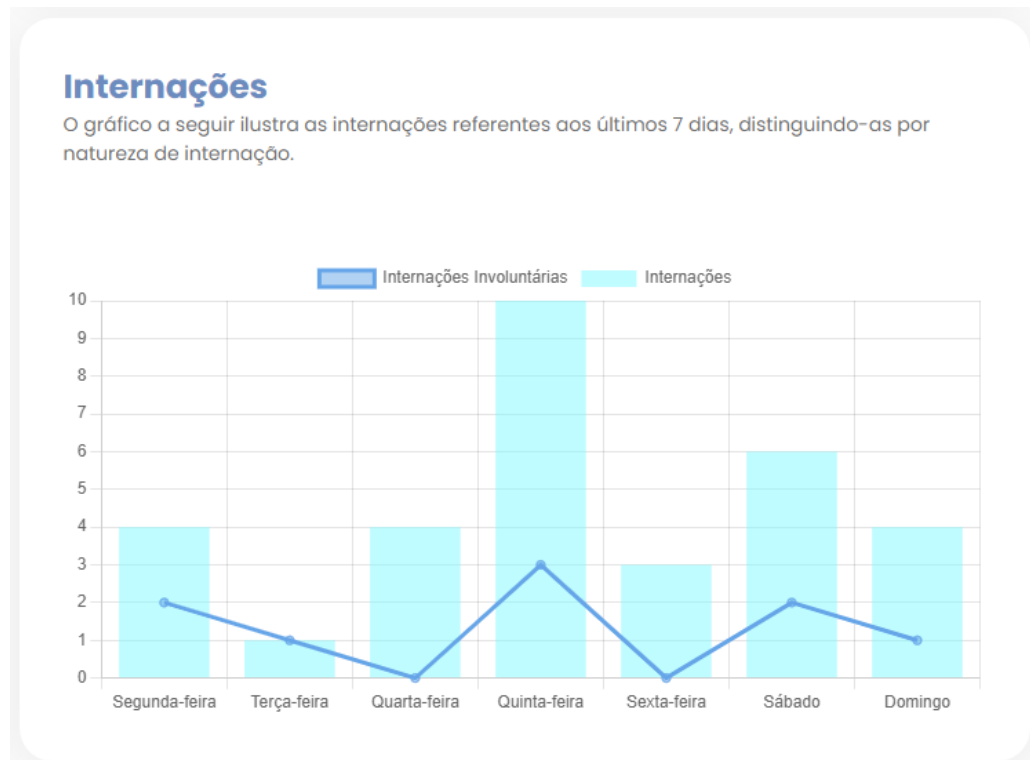
Fonte: Elaborado pelo autor

**Figura 8 - Gráfico de emergências presente na página Home**



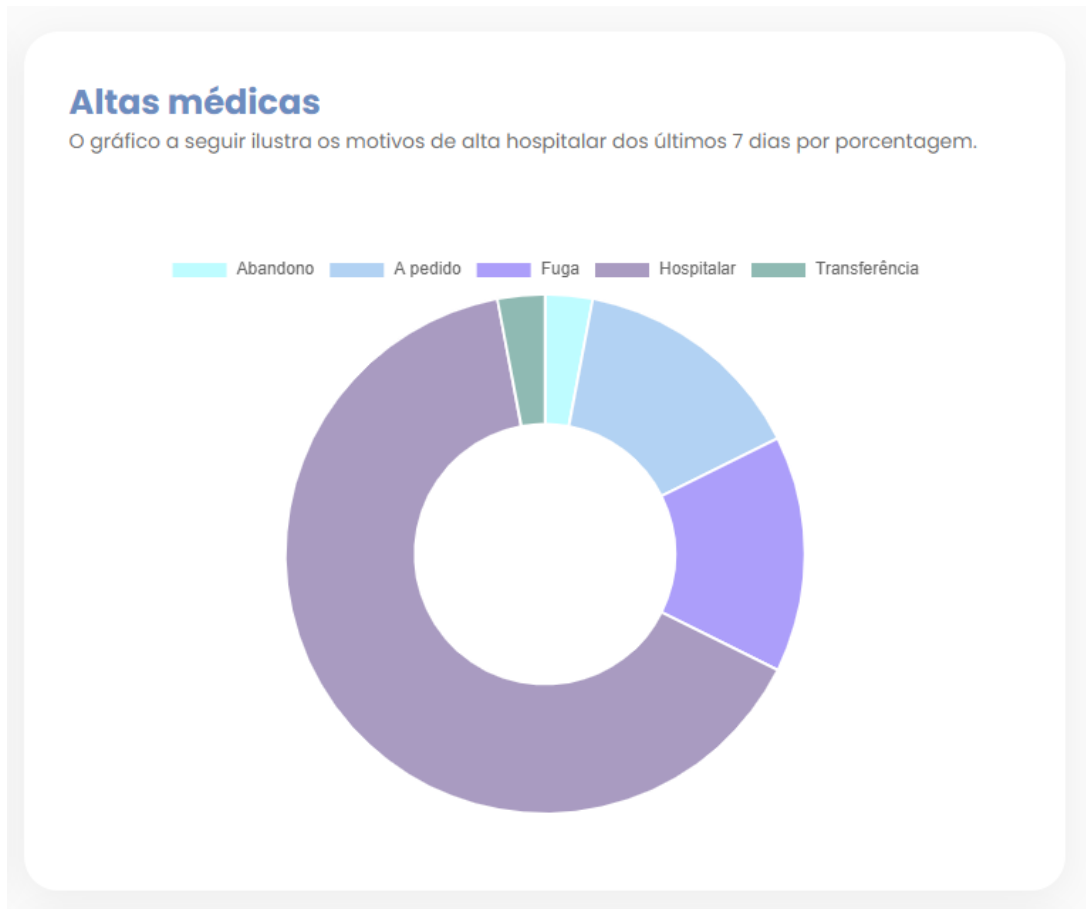
Fonte: Elaborado pelo autor.

**Figura 9 - Gráfico de Internações presente na página Home**



Fonte: Elaborado pelo autor.

**Figura 10 - Gráfico de Altas Médicas presente na página Home**

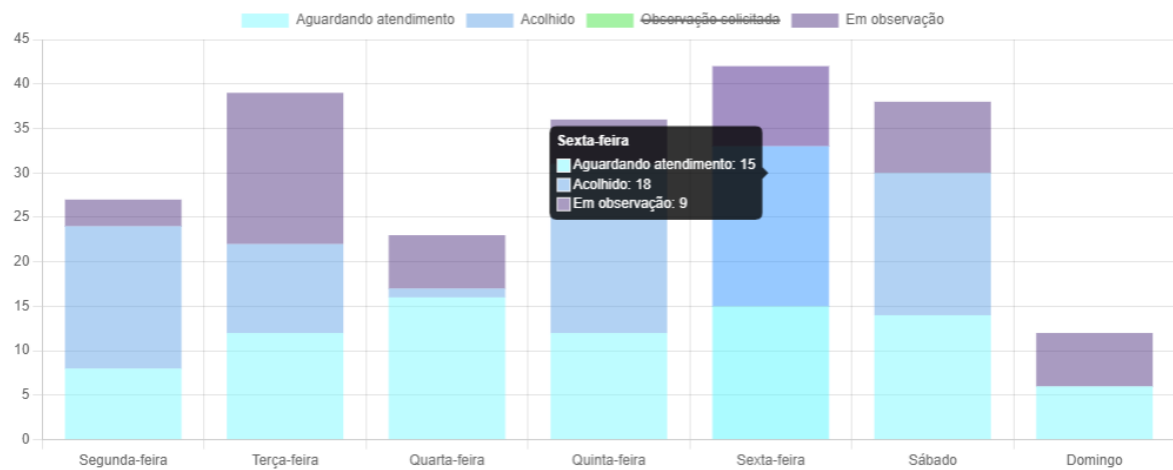


Fonte – Elaborado pelo autor.

**Figura 11 - Recursos de manipulação**

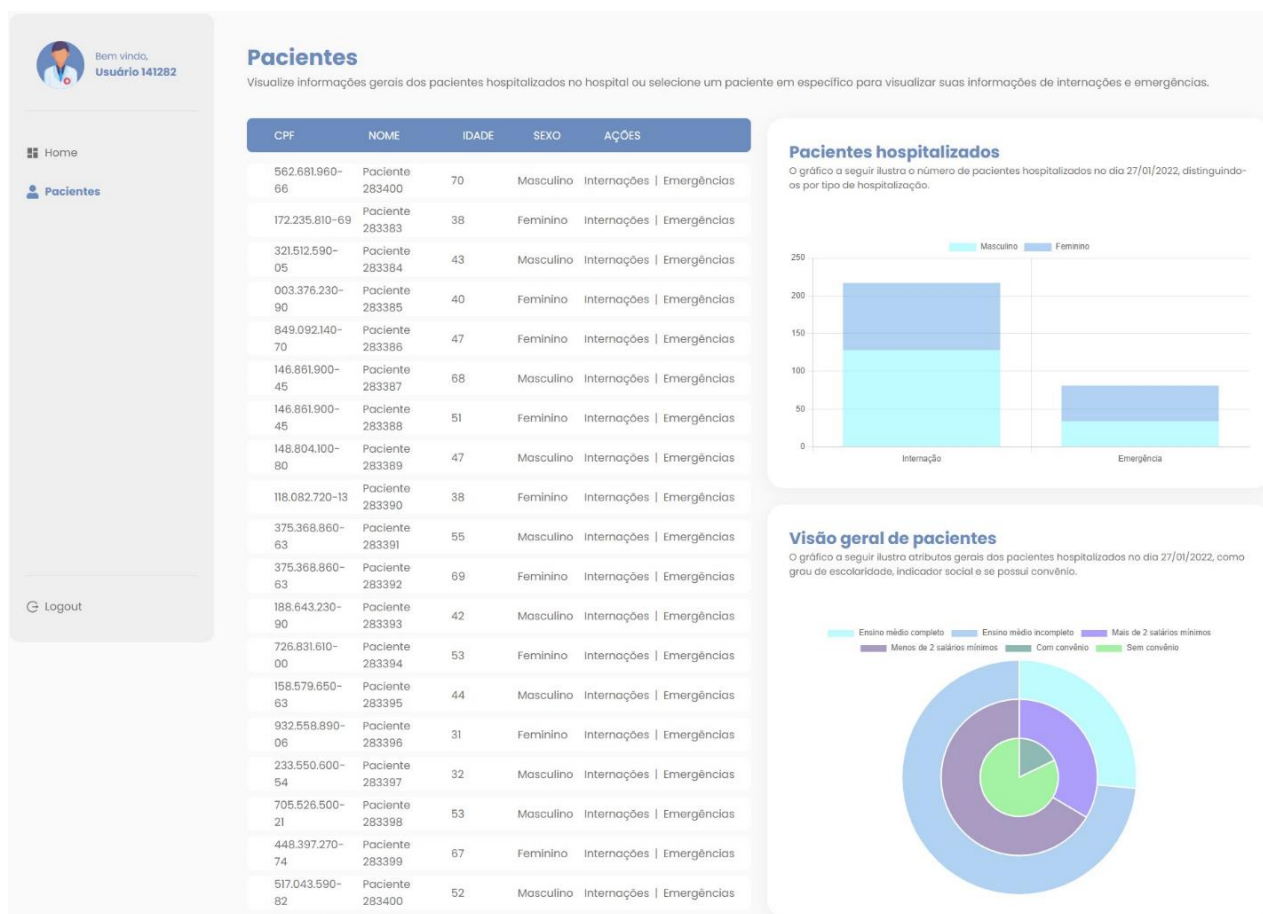
**Emergências**

O gráfico a seguir ilustra os atendimentos de emergência referentes aos últimos 7 dias, distinguindo-os por estado de atendimento.



Fonte – Elaborado pelo autor.

Figura 12 - Tela de pacientes



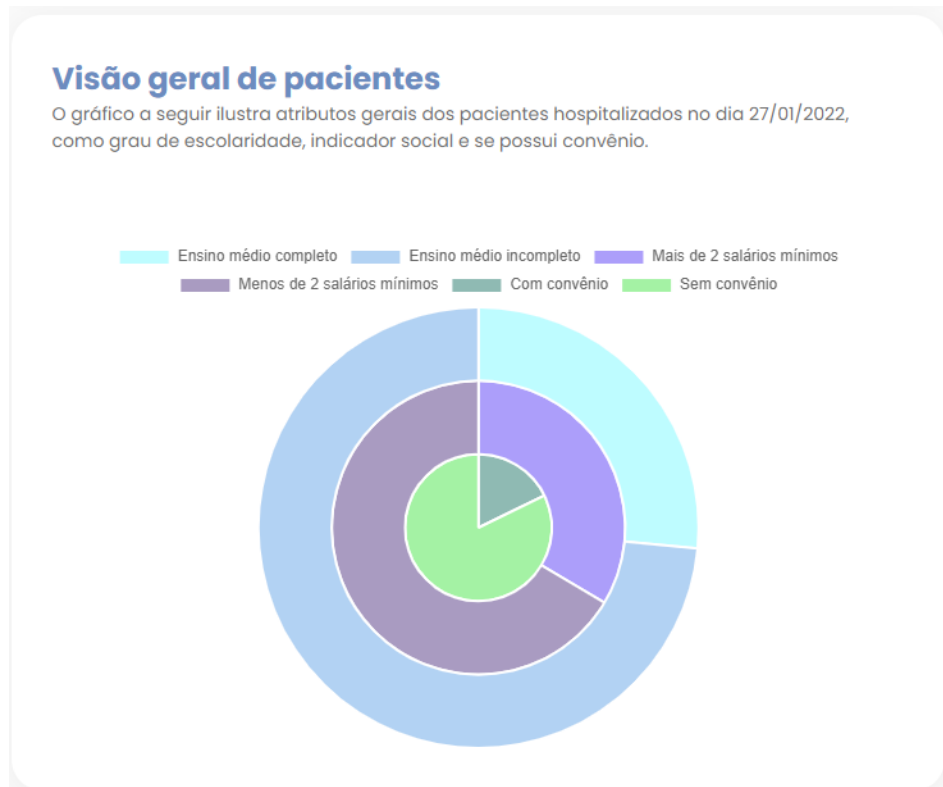
Fonte – Elaborado pelo autor

### 3.5.3 Pacientes

A página de pacientes está ilustrada na Figura 12 e pode ser dividida em duas seções. A primeira, presente à esquerda, trata-se da listagem de pacientes hospitalizados naquele momento. Tal listagem exibe informações básicas de identificação do paciente, como CPF, nome, idade e sexo. Além disso, para cada um dos pacientes presentes na lista, é possível acessar por meio da coluna AÇÕES seus dados de internações e seus atendimentos de emergência.

À direita da página, encontram-se gráficos que apresentam informações gerais dos pacientes hospitalizados no momento, como o número de internações e atendimentos de emergência, categorizados por sexo do paciente. Há, ainda, o gráfico de visão geral, que contém informações como grau de escolaridade, indicador social e se o paciente em questão trata-se de um paciente de convênio ou da rede pública.

**Figura 13 - Gráfico de Visão geral de pacientes presente na página Pacientes**



Fonte – Elaborado pelo autor

**Figura 14 - Gráfico de Pacientes hospitalizados presente na página de Pacientes**



Fonte – Elaborado pelo autor

### 3.5.4 Internações do paciente

A página de internações do paciente exibe informações referentes a cada uma das internações de um paciente dentro do hospital. À esquerda, pode-se visualizar a listagem de internações com informações de identificação, tais como, médico responsável, data de entrada, tipo do paciente e natureza da internação.

Ao clicar em uma das internações presentes na lista, é exibido à esquerda da página, um gráfico que exibe os diferentes CIDs contidos na internação selecionada. Tal informação consta nas notas clínicas do paciente referentes à internação em questão, de modo que estas são resultantes do processamento de dados não estruturados. A página de internações do paciente está ilustrada na Figura 15.

**Figura 15 - Página de Internações**



Fonte – Elaborado pelo autor.



**Figura 16 - Gráfico de CIDs da Internação presente na página de Internação**



Fonte – Elaborado pelo autor.

### 3.5.5 Atendimentos de emergência do paciente

A página de atendimentos de emergência de determinado paciente é semelhante à página de internações do paciente. Esta é igualmente separada em dois principais componentes. À esquerda, encontra-se uma lista com os atendimentos de emergência do paciente. Analogamente ao componente de listagem de página de internações, este contém informações básicas de identificação do atendimento de emergência, como médico de observação, data de entrada e tipo do paciente.

Além disso, ao clicar em um dos atendimentos de emergência listados, um gráfico é renderizado a direita da página, exibindo informações relativas aos CIDs contidos nas notas clínicas referentes ao atendimento de emergência selecionado. Da mesma forma, tais informações são retiradas de fontes de dados não estruturados e seriam de difícil acesso

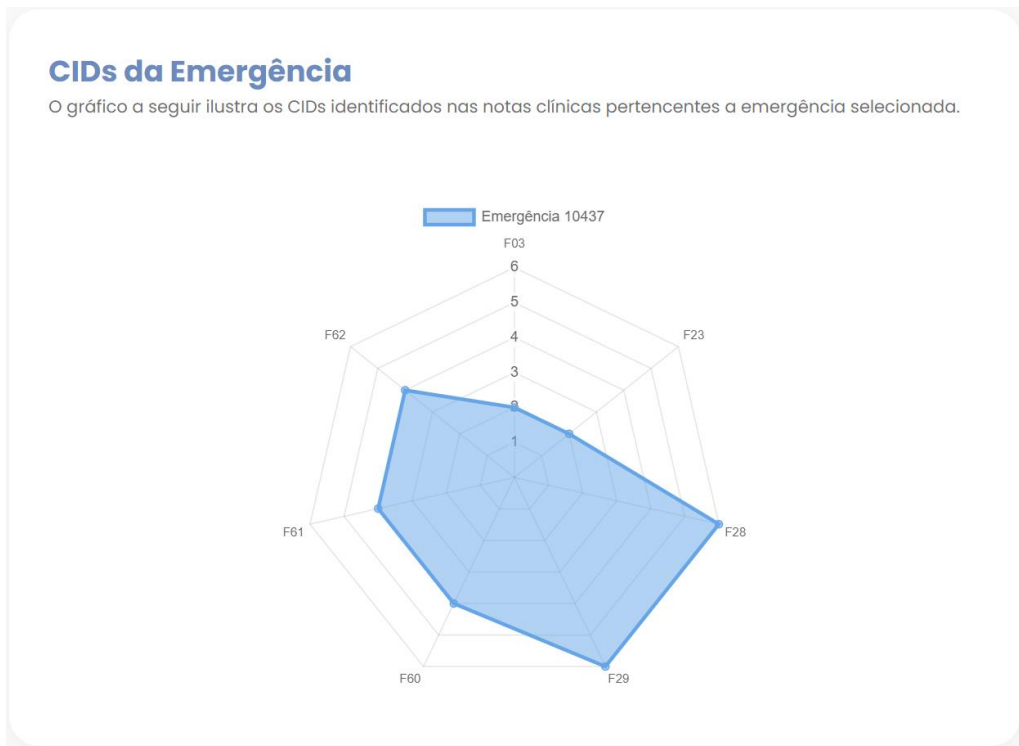
mediante processamento manual das notas clínicas. Na Figura 17 está ilustrada a página de atendimentos de emergência do paciente.

**Figura 17 - Tela de atendimentos de emergência de um paciente**



Fonte – Elaborado pelo autor.

**Figura 18 - Gráfico de CID's de Emergência presente na página de Emergências**



Fonte – Elaborado pelo autor.

## 3.6 Considerações finais

Neste capítulo, inicialmente foi apresentada a arquitetura utilizada para o desenvolvimento da aplicação. A mesma é baseada em microsserviços que possuem diferentes responsabilidades, como acesso a dados estruturados e não estruturados. Tais componentes comunicam-se por meio de um sistema de filas de mensagens assíncrono, o que fornece maior tolerância à falha em um dos lados da comunicação. Por fim, ambos os microsserviços são orquestrados por um controlador, que recebe as chamadas da interface e repassa para o componente responsável por oferecer tais recursos.

Em seguida, foi apresentado o ambiente de visualização desenvolvido, que consiste em uma plataforma *web* capaz de apresentar de forma visual dados de diferentes fontes e tipos — estruturados e não estruturados. Foi dito, ainda, que as visualizações presentes no ambiente contam com recursos de manipulação de dados, importantes para o processo de descobrimento de padrões.

Portanto, neste trabalho foi desenvolvido um ambiente de visualização capaz de trabalhar com as diferentes formas em que se apresentam os dados originados pelo setor da saúde, de modo que estes podem ser estruturados ou não estruturados. Além disso, tal ambiente foi desenvolvido por meio de uma arquitetura baseada em microsserviços, estratégia que garante maior escalabilidade e reusabilidade de código, fatores que apesar de importantes para *softwares* da área da saúde, são fortemente limitados na maioria das aplicações do ramo, devido ao projeto de arquitetura utilizado nestes sistemas.

# Capítulo 4

## Avaliação experimental

### 4.1 Considerações iniciais

Este capítulo tem como objetivo expor a metodologia de experimentação utilizada para validar o ambiente de visualização proposto. Em seguida, os resultados dos testes são exibidos e discutidos.

### 4.2 Metodologia de experimentação

Realizar experimentos com o intuito de validar ambientes de visualização pode se tornar uma atividade significativamente complexa, uma vez que estas possuem muitos componentes distintos entre si, de diferentes níveis de complexidade. Tais componentes incluem raciocínio analítico do usuário, as diferentes representações visuais, técnicas de interface humano-computador, técnicas de transformações de dados, entre outras (PLAISANT; GRINSTEIN; SCHOLTZ, 2009).

Desta forma, o presente ambiente de visualização foi avaliado a partir de uma abordagem heurística bem consolidada na literatura científica, uma vez que tais métodos são comuns em estudos acerca da usabilidade de sistemas de informações. O procedimento utilizado para avaliação consiste em uma combinação de diferentes princípios heurísticos que visam uma forma padrão de avaliar sistemas de visualização empregados em ambientes de saúde (DOWDING; MERRILL, 2018).

Diante do exposto, o método de experimentação do ambiente de visualização foi realizado mediante a aplicação de um teste de uso, por meio da aplicação de um questionário para um conjunto de usuários convidados. Tal questionário foi embasado na coleção de princípios de usabilidade e nos fatores a estes associados, segundo critérios apontados por

Dowding e Marrill (2018) e encontra-se disponível no apêndice A. Nele, os seguintes princípios foram avaliados:

- a) Visibilidade do status do sistema – Avalia a capacidade do sistema de manter o usuário informado acerca de todos os acontecimentos em tempo de uso, por meio de um *feedback* apropriado e dentro de um tempo razoável.
- b) Correspondência entre o sistema e o mundo real – Avalia a capacidade do sistema em comunicar-se no mesmo idioma de seu usuário, por meio de palavras, frases e conceitos familiares ao mesmo, no lugar de termos orientados pelo sistema, de modo que, para tal, é de suma importância que o sistema avaliado utilize convenções bem definidas e exiba informações de maneira natural e lógica.
- c) Controle e liberdade do usuário – Avalia a capacidade do sistema de fornecer liberdade para que seus usuários possam selecionar e sequenciar tarefas, quando apropriado. Para tal, os usuários necessitam de meios para interromper o estado atual de uma atividade em andamento, de modo que facilmente possam iniciar uma segunda atividade.
- d) Consistências e padrões – Avalia a capacidade do sistema de padronizar palavras, ícones, termos, expressões, entre outros, de modo que um mesmo conceito deve sempre ter o mesmo significado.
- e) Reconhecimento no lugar de recordação – Capacidade do sistema de apresentar objetos, ações e opções de forma clara, de modo que o usuário não necessite lembrar-se de instruções ou informações para operar determinado módulo.
- f) Flexibilidade e eficiência de uso – Avalia a capacidade do sistema de oferecer ao usuário uma gama de opções no que diz respeito a localização de conteúdo, de modo que o usuário se torna capaz de atingir seus objetivos de forma mais eficiente.
- g) Design estético e minimalista – Avalia a capacidade do sistema em apresentar apenas informações relevantes para seu usuário, ou seja, componentes do sistema não devem conter informações irrelevantes ou raramente necessárias, uma vez que cada unidade extra de informação compete com as unidades relevantes.

- h) Organização espacial – Avalia a capacidade do sistema de fazer com que seus principais componentes sejam facilmente localizados, isto é, a forma como seus diversos componentes estão distribuídos ao longo do sistema.
- i) Codificação de informação – Avalia se o sistema utiliza símbolos ou representações que auxiliem no processo de percepção de seus usuários.
- j) Orientação e ajuda – Avalia se o sistema oferece o suporte necessário para o usuário, ou seja, a capacidade do sistema de ajudar e orientar seu usuário.

Além disso, é válido ressaltar que para a realização dos testes de uso, o ambiente de visualização foi implantado em um servidor, de modo que os usuários convidados a participar dos testes foram capazes de acessá-lo por meio da *internet*. Por fim, o questionário para avaliação do sistema foi aplicado por meio do software *Google Forms* (Google, 2018). Este, consiste em uma aplicação web para produção, compartilhamento e análise de formulários, de modo que tais formulários se mantêm disponíveis para acesso online.

### 4.3 Resultados

A Tabela 1 a seguir exibe a classificação de cada um dos princípios heurísticos avaliados pela pesquisa com os usuários convidados:

**Tabela 1 – Resultados do teste de uso**

Princípio	Pontuação máxima	Pontuação média	Resultados
Visibilidade do status do sistema	6	4,82	80,33%
Correspondência entre o sistema e o mundo real	5	3,82	76,40%
Controle e liberdade do usuário	5	2,88	57,60%
Consistências e padrões	6	5,06	84,33%
Reconhecimento no lugar de recordação	5	4,53	90,60%
Flexibilidade e eficiência de uso	4	2,12	53%
Design estético e minimalista	5	4,24	84,80%
Organização espacial	3	2,70	90%
Codificação de informação	2	1,59	79,50%
Orientação e ajuda	3	1,59	53%
Total	44	33,35	75,80%

**Fonte** – Elaborado pelo autor.

Para cada princípio da heurística utilizada há uma pontuação máxima a ser alcançada. Cada um destes pontos corresponde ao número de fatores que o usuário considerou ao avaliar o sistema. Em outras palavras, cada um dos pontos corresponde a uma das questões contidas no questionário. Portanto, pode-se utilizar como exemplo o princípio heurístico visibilidade de status do sistema. Este possui uma pontuação máxima de seis pontos, equivalentes às seis questões presentes no questionário, de tal modo que uma vez que o convidado avaliar como verdadeiro o fator questionado (ou seja, responder sim à questão do formulário), atribui-se o ponto deste fator ao princípio avaliado. Desta forma, foi calculada a pontuação média para cada princípio abordado no questionário.

Uma vez avaliados os resultados obtidos no teste de uso, destacam-se como as menores pontuações médias os princípios: controle e liberdade do usuário, flexibilidade e eficiência de uso e orientação e ajuda. Portanto, pode-se destacá-los como pontos em que o ambiente de visualização não atingiu um resultado satisfatório na visão dos usuários convidados a participar do teste. Além disso, é possível avaliar ainda os fatores individuais que mais influenciaram para que tais quesitos não alcançassem pontuações melhores.

Para o princípio de usabilidade de controle e liberdade do usuário, a falta de uma função para desfazer ações bem como a falta de recursos para que o usuário possa acessar todas as páginas da aplicação ao longo do sistema impactaram negativamente em sua pontuação. A Tabela 2 a seguir apresenta os resultados do questionário sobre controle e liberdade do usuário.

**Tabela 2 – Resultados do fator Controle e liberdade do usuário**

Questão	Número de respostas “Sim”	Número de respostas “Não”
1	15	2
2	4	13
3	2	15
4	12	5
5	16	1

**Fonte** – Elaborado pelo autor

Em seguida, pode-se realizar a mesma análise para o princípio heurístico flexibilidade e eficiência de uso, que obteve, em média, apenas 2,12 pontos dos 4 pontos máximos. Isto se dá, pois, o ambiente de visualização não oferece recursos para que funcionalidades como, navegação por meio de apontadores e acesso a itens de determinado menu mediante atalhos de

teclado estivessem disponíveis para o usuário final. A Tabela 3 a seguir apresenta os resultados do questionário para o quesito flexibilidade e eficiência de uso.

**Tabela 3 – Resultados do fator Flexibilidade e eficiência de uso**

Questão	Número de respostas “Sim”	Número de respostas “Não”
1	14	3
2	12	5
3	4	13
4	6	11

**Fonte** – Elaborado pelo autor

Pode-se citar ainda o quesito de orientação e ajuda, que obteve 1,59 dos 3 pontos disponíveis. Uma vez que os resultados do questionário relativo ao princípio heurístico em questão são avaliados, torna-se notável como, novamente, a falta de recursos para que o usuário possa retomar em suas ações em conjunto com a falta de um feedback visual que informe o caminho seguido para chegar em determinada localização do sistema tiveram forte impacto na avaliação do quesito testado. Os resultados referentes a avaliação do princípio heurístico em questão encontram-se na Tabela 4 a seguir.

**Tabela 4 – Resultados do fator Orientação e ajuda**

Questão	Número de respostas “Sim”	Número de respostas “Não”
1	16	1
2	4	13
3	7	10

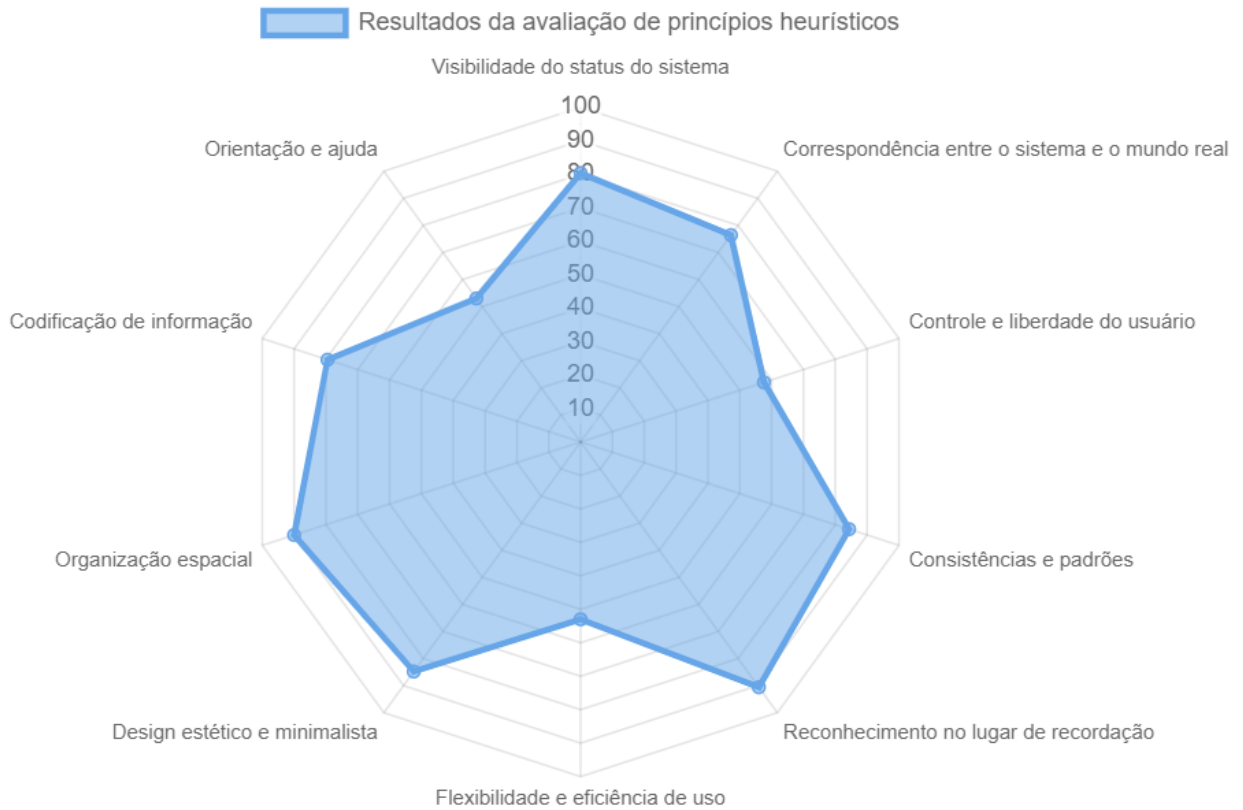
**Fonte** – Elaborado pelo autor

Por fim, todos os demais princípios heurísticos avaliados obtiveram uma pontuação média superior a 70%, o que indica que o ambiente de visualização foi eficaz em atender a maior parte dos fatores de usabilidade expostos no questionário de testes. Portanto, uma vez avaliado de forma geral, o ambiente de visualização desenvolvido foi avaliado com 33.35 pontos dos 44 disponíveis, percentual de 75,80% dos pontos. Disso, pode-se concluir que como um todo, o ambiente de visualização desenvolvido possui um valor agregado e que os recursos disponíveis pelo mesmo contribuem com a diminuição da chance de erros de interpretação ou, ainda, significados gerados por visualizações inadequadas. Na Figura 11 a seguir está ilustrada



a avaliação de todos os princípios heurísticos discutidos para o ambiente de visualização desenvolvido.

**Figura 19 – Resultados da avaliação dos princípios heurísticos**



Fonte – Elaborado pelo autor.

## 4.4 Considerações finais

Por meio dos resultados obtidos, pode-se concluir que o ambiente de visualização desenvolvido é efetivo e possui fatores que contribuem para aprimorar a usabilidade do usuário. Fator que beneficia o processo de extração de conhecimento por meio das visualizações nele contidas. Contudo, existem ainda fatores a serem aperfeiçoados com o intuito de maximizar as métricas de usabilidade do sistema e conseqüentemente, causar um impacto positivo também no processo de descobrimento de padrões por meio das visualizações.

# Capítulo 5

## Conclusões

### 5.1 Considerações iniciais

Com o surgimento do Prontuário Eletrônico do Paciente (PEP), o setor da saúde passou a contar com recursos para geração e compartilhamento de dados em tempo real (MENACHEMI; COLLUM, 2011). Como uma forma eficiente de lidar com este grande volume de dados, a extração de conhecimento por meio de técnicas visuais apresenta-se como uma alternativa valiosa, uma vez que esta cumpre de forma satisfatória a tarefa de integrar o usuário ao processo de exploração dos dados. Porém, para que tais dados possam ser utilizados de forma eficaz, é necessário que seja possível lidar com suas diferentes características, sejam estes estruturados ou não estruturados. Além disso, os sistemas de visualização presentes na área da saúde encontram dificuldades com relação à arquitetura em que são desenvolvidos, visto que esta normalmente apresenta problemas graves, como: ausência de práticas de reuso de software, limitação de escalabilidade, maiores custos, ausência de recursos para garantir confiabilidade e disponibilidade, entre outros (DA SILVA et al, 2019).

Diante deste cenário, o objetivo deste trabalho foi desenvolver um ambiente de visualização de dados eficiente em lidar com os diferentes tipos de dados presentes no domínio da saúde, sejam estes estruturados ou não estruturados. Além disso, o intuito foi de prover, ainda, um sistema baseado em uma arquitetura de microsserviços capaz de ser implantado em diferentes ambientes de

*cloud*. Desta forma, garante-se um sistema eficiente na exploração visual de dados estruturados e não estruturados que está preparado para escalar, com recursos que garantem a confiabilidade e disponibilidade do mesmo, além de utilizar de práticas de reuso. Tais fatores impactam positivamente em seu custo.

Para que o objetivo fosse alcançado, foi desenvolvido um ambiente de visualização de dados da saúde baseado em uma arquitetura de microsserviços, de modo que as diferentes fontes de dados fossem separadas em serviços diferentes. Serviços de acesso a dados estruturados foram responsáveis por acessar a base de dados do hospital parceiro do Grupo de Banco de Dados (GBD) e encaminhá-los para o ambiente de visualização. Por sua vez, serviços de acesso a dados não estruturados foram integrados com um armazém de dados, que armazena registros originalmente contidos em notas clínicas. A comunicação entre os serviços foi realizada por meio de filas de mensagens, a fim de garantir maior performance e tolerância a falhas.

Com o ambiente de visualização desenvolvido, este foi avaliado com base em um método heurístico, com o intuito de simplificar o impacto de diversos fatores complexos e muitas vezes subjetivos que impactam na avaliação do ambiente de visualização, como raciocínio analítico do usuário, as diferentes representações visuais, técnicas de interface humano-computador, técnicas de transformações de dados, entre outras (PLAISANT; GRINSTEIN; SCHOLTZ, 2009).

## **5.2 Contribuições científicas**

Os ambientes de visualização presentes na literatura falham em lidar com as diferentes formas em que se apresentam os dados presentes no setor da saúde por meio de arquiteturas que visam reduzir os custos e possibilitar maior escalabilidade, tolerância a falhas, alta disponibilidade e confiabilidade (DA SILVA et al, 2019). Portanto, a contribuição científica deste trabalho é o desenvolvimento de um ambiente de visualização de dados capaz de lidar com dados estruturados e não estruturados, baseado em uma arquitetura de microsserviços que possibilita maiores níveis de escalabilidade, confiabilidade, disponibilidade e, ainda, minimização dos custos do sistema.

No Quadro 3 são apresentadas as comparações realizadas entre este trabalho e os demais trabalhos da área, presentes na literatura científica e abordados na seção 2.8. Evidentemente, apenas o trabalho desenvolvido é eficiente em todos os aspectos presentes no quadro.

Quadro 3 - Comparação com trabalhos correlatos e este trabalho

	(SHABAN-NEJAD et al., 2017)	(RUAN et al., 2018)	(TANG et al., 2019)	(BARROCA FILHO et al., 2020)	Trabalho desenvolvido
Capaz de visualizar dados estruturados	✓	✗	✗	✓	✓
Capaz de visualizar dados não estruturados	✓	✓	✓	✗	✓
Utiliza recursos de <i>Cloud Computing</i>	✗	✗	✗	✓	✓
Utiliza arquitetura baseada em microsserviços	✗	✗	✗	✗	✓

Fonte: Elaborado pelo autor.

### 5.3 Trabalhos futuros

Após a análise dos resultados obtidos na fase de testes, observou-se a possibilidade de realizar melhorias no ambiente de visualização desenvolvido. Tais melhorias são citadas a seguir:

- Adaptar o ambiente de visualização para diferentes dispositivos, como smartphones e tablets, de modo a aperfeiçoar a experiência dos usuários ao utilizar o ambiente de visualização em outros dispositivos;
- Incorporar camadas de *cache* nas chamadas aos microsserviços de acesso a dados estruturados e não estruturados, com o objetivo de melhorar a performance da aplicação;
- Adicionar outros tipos de visualização, de forma a ampliar o leque de possibilidades durante o processo de exploração de dados;
- Criação de um *framework* capaz de encapsular os itens comuns da arquitetura, de modo a simplificar o processo de implementação;

- e) Adaptar o ambiente de desenvolvimento para a utilização de *microfrontends*, a fim de torná-lo mais flexível e adaptável.

## 5.4 Considerações finais

Uma vez que o foco do trabalho era a construção do ambiente de visualização de modo que este fosse capaz de lidar com os diferentes tipos de dados da saúde e, ainda, ser eficiente e estar preparado para lidar com requisitos frequentemente exigidos no setor da saúde, como capacidade de alta escalabilidade, tolerância a falhas, alta disponibilidade, alta confiabilidade e, por fim, baixos custos, pode-se concluir, portanto, que as expectativas foram atendidas, visto que o ambiente desenvolvido foi implementado baseado em uma arquitetura de microsserviços que suporta tais requisitos e obteve notas satisfatórias nos testes realizados com os usuários convidados.

# REFERÊNCIAS

ANDREU-PEREZ, Javier; POON Carmen; MARRIFIELD Robert; WONG Stephen; YANG Guang-Zhong. **IEEE journal of biomedical and health informatics**, v. 19, n. 4, p. 1193-1208, 2015.

APACHE OPENNLP, OpenNLP. Disponível em: <https://opennlp.apache.org/>. Acesso em 24 de julho de 2021.

BARROCA FILHO, Itamir et al. Development of a Health Dashboard for an Electronic Health Record System. **In: 2020 20th International Conference on Computational Science and Its Applications (ICCSA)**. IEEE, 2020. p. 16-22.

BRASIL. Resolução CFM nº 1.638/2002. Define prontuário médico e torna obrigatória a criação da Comissão de Revisão de Prontuários nas instituições de saúde. **Diário Oficial da União**. 9 ago 2002, Seção I, p.184-5.

CHEN, Wei; GUO, Fangzhou; WANG, Fei-Yue. A survey of traffic data visualization. **IEEE Transactions on Intelligent Transportation Systems**, v. 16, n. 6, p. 2970-2984, 2015.

CHEN, M; MAO, S; LIU, Y. Big data: A survey. **Mobile networks and applications**, v. 19, n. 2, p. 171-209, 2014.

COWIE, Martin R.; BLOMSTER, Juuso L.; CURTIS, Lesley H.; DUCLAUX, Sylvie. FORD, Ian; FRITZ, Fleur; GOLDMAN, Samantha; JANMOHAMED, Salim; KREUZER, Jorg; LEENAY, Mark; MICHEL, Alexander; ONG, Seleen; PELL, Jill P.; SOUTHWORTH Ross Mary; STOUGH, Gattis Wendy; THOENES Martin; ZANNAD Faiez; ZALEWSKI Andrew. **Electronic health records to facilitate clinical research. Clinical Research in Cardiology**, v. 106, n. 1, p. 1-9, 2017.

DA SILVA, Marcio Alexandre Pereira; TIMES, Valéria Cesária; DE ARAÚJO, André Magno Costa; DA SILVA, Paulo Caetano. A Microservice-Based Approach for Increasing Software Reusability in Health Applications. In: **2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)**. IEEE, 2019. p. 1-8.

DOWDING, Dawn; MERRILL, Jacqueline A. The development of heuristics for evaluation of dashboard visualizations. **Applied clinical informatics**, v. 9, n. 03, p. 511-518, 2018.

ELASTIC, Elasticsearch. Disponível em: <https://www.elastic.co/pt/elasticsearch/>. Acesso em 25 de julho de 2021a.

ELASTIC, Logstash. Disponível em: <https://www.elastic.co/pt/logstash/>. Acesso em 25 de julho de 2021b.

ELASTIC, Kibana. Disponível em: <https://www.elastic.co/pt/kibana/>. Acesso em 25 de julho de 2021c.

GOOGLE. **Google Forms**, 2021. Disponível em <<https://www.google.com/forms>>. Acesso em: 04 de novembro. 2021.

HEWAGAMAGE, K. Priyantha; HIRAKAWA, Masahito; ICHIKAWA, Tadao. Interactive visualization of spatiotemporal patterns using spirals on a geographical map. In: **Proceedings 1999 IEEE Symposium on Visual Languages**. IEEE, 1999. p. 296-303.

JONES, Michael B.; BRADLEY, John; SAKIMURA, Nat. **RFC 7519: JSON Web Token (JWT)**, 2015.

JOSEPH, Renien John. Single page application and canvas drawing. **arXiv preprint arXiv:1502.03530**, 2015.

JWT.Auth0. Disponível em: <https://jwt.io/>. Acessado em: 20 de outubro de 2021.

KEIM, Daniel A. Information visualization and visual data mining. **IEEE transactions on Visualization and Computer Graphics**, v. 8, n. 1, p. 1-8, 2002.

KOHATSU, Neal D.; ROBINSON, Jennifer G.; TORNER, James C. Evidence-based public health: an evolving concept. **American journal of preventive medicine**, v. 27, n. 5, p. 417-421, 2004.

KUO, Mu-Hsing. Opportunities and challenges of cloud computing to improve health care services. **Journal of medical Internet research**, v. 13, n. 3, p. e67, 2011.

LANEY, Doug. 3D data management: Controlling data volume, velocity and variety. **META group research note**, v. 6, n. 70, p. 1, 2001.

MEIJER, Erik. The world according to LINQ. **Communications of the ACM**, v. 54, n. 10, p. 45-51, 2011.

MENACHEMI, Nir; COLLUM, Taleah H. Benefits and drawbacks of electronic health record systems. **Risk management and healthcare policy**, v. 4, p. 47, 2011.

MESBAH, Ali; VAN DEURSEN, Arie. An architectural style for Ajax. In: **2007 Working IEEE/IFIP Conference on Software Architecture (WICSA'07)**. IEEE, 2007. p. 9-9.

METAMAP, MetaMap. Disponível em: <https://metamap.nlm.nih.gov/>. Acesso em 24 de julho de 2021.

NATIONAL LIBRARY OF MEDICINE (NLM), Unified Medical Language System. Disponível em: <https://www.nlm.nih.gov/research/umls/index.html>. Acesso em: 24 de julho de 2021.

PLAISANT, Catherine; GRINSTEIN, Georges; SCHOLTZ, Jean. Visual-analytics evaluation. **IEEE Computer Graphics and Applications**, v. 29, n. 03, p. 16-17, 2009.



REDDY, Chandan K.; AGGARWAL, Charu C. (Ed.). **Healthcare data analytics**. CRC Press, 2015.

RIMAL, Bhaskar Prasad; CHOI, Eunmi; LUMB, Ian. A taxonomy and survey of cloud computing systems. In: **2009 Fifth International Joint Conference on INC, IMS and IDC**. Ieee, 2009. p. 44-51

RUAN, Wei et al. Pictorial visualization of EMR summary interface and medical information extraction of clinical notes. In: **2018 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)**. IEEE, 2018. p. 1-6.

SHABAN-NEJAD, Arash et al. PopHR: a knowledge-based platform to support integration, analysis, and visualization of population health data. **Annals of the New York Academy of Sciences**, v. 1387, n. 1, p. 44-53, 2017.

SUN, W., CAI, Z., LI, Y., LIU, F., FANG, S., & WANG, G. Data processing and text mining technologies on electronic medical records: a review. **Journal of healthcare engineering**, v. 2018, 2018.

TANG, Chunlei et al. A temporal visualization of chronic obstructive pulmonary disease progression using deep learning and unstructured clinical notes. **BMC medical informatics and decision making**, v. 19, n. 8, p. 1-9, 2019.

TAYEFI, Maryam; NGO, Phuong; CHOMUTARE, Taridzo; DALIANIS, Hercules; SALVI, Elisa; BUDRIONIS, Andrius; GODTLIEBSEN, Fred. Challenges and opportunities beyond structured data in analysis of electronic health records. **Wiley Interdisciplinary Reviews: Computational Statistics**, p. e1549, 2021.

VINOSKI, Steve. Advanced message queuing protocol. **IEEE Internet Computing**, v. 10, n. 6, p. 87-89, 2006.

WORKNEH, Fekadu; ADEM, Ahmed; PRADHAN, Roshni. Understanding cloud based health care service with its benefits. In: **2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)**. IEEE, 2018. p. 102-106.

ZIKOPOULOS, Paul; EATON, Chris. **Understanding big data: Analytics for enterprise class hadoop and streaming data**. McGraw-Hill Osborne Media, 2011.

# **Apêndice A – Questionário de Avaliação do Ambiente de Visualização de Dados**

## **Questionário sobre visibilidade e status do sistema**

**1. O ambiente de visualização possui em cada uma de suas telas títulos ou cabeçalhos que descrevam seu conteúdo?**

a) Sim

b) Não

**2. O ambiente de visualização possui um esquema de design de ícones consistente e um respectivo tratamento estilístico?**

a) Sim

b) Não

**3. O ambiente de visualização oferece feedback visual em menus ou caixas de diálogo sobre quais opções podem ser selecionadas?**

a) Sim

b) Não

**4. O ambiente de visualização exibe uma indicação clara da localização atual?**

a) Sim

b) Não

**5. O ambiente de visualização possui uma terminologia de nomenclatura de menu consistente com o domínio da saúde?**

- a) Sim
- b) Não

**6. O ambiente de visualização oferece recursos para que o usuário identifique o estado atual do sistema e as alternativas de ação?**

- a) Sim
- b) Não

## **Questionário sobre correspondência entre o sistema e o mundo real**

**1. O ambiente de visualização possui ícones concretos e familiares?**

- a) Sim
- b) Não

**2. O ambiente de visualização possui seções e subseções onde os títulos estão ordenados de maneira lógica?**

- a) Sim
- b) Não

**3. O ambiente de visualização possui menus que possuem uma sequência natural em seus itens?**

- a) Sim
- b) Não

**4. O ambiente de visualização possui cores correspondentes às expectativas comuns sobre os códigos de cores?**

- a) Sim
- b) Não

**5. O ambiente de visualização usa de palavras e conceitos familiares aos usuários?**

- a) Sim
- b) Não

## **Questionário sobre controle e liberdade do usuário**

**1. O ambiente de visualização disponibiliza uma forma clara de sair de cada uma das telas?**

- c) Sim
- d) Não

**2. O ambiente de visualização oferece recursos para que todas as telas sejam acessíveis a partir de qualquer localização do sistema?**

- a) Sim
- b) Não

**3. O ambiente de visualização oferece funções de desfazer?**

- a) Sim
- b) Não

**4. O ambiente de visualização oferece recursos para que o usuário possa clicar nos itens do menu com o mouse ou usar um atalho do teclado?**

- a) Sim
- b) Não

**5. O ambiente de visualização oferece recursos para que o usuário avance e retroceda entre as telas facilmente?**

- a) Sim
- b) Não

## **Questionário sobre padrões e consistência**

**1. O ambiente de visualização possui um padrão de formatação seguido de forma consistente em todas as telas do sistema?**

- a) Sim
- b) Não

**2. O ambiente de visualização possui título em cada uma das janelas?**

- a) Sim
- b) Não

**3. O ambiente de visualização possui pistas visuais salientes para identificar a tela ativa?**

- a) Sim
- b) Não

**4. O ambiente de visualização não possui componentes com alto número de cores (de quatro a sete) unidas no mesmo espectro visível?**

a) Sim

b) Não

**5. O ambiente de visualização possui nomes consistentes no que diz respeito a estilo gramatical e terminologia?**

a) Sim

b) Não

**6. O ambiente de visualização possui listas de opções de menu apresentadas verticalmente?**

a) Sim

b) Não

## **Questionário sobre reconhecimento no lugar de recordação**

**1. O ambiente de visualização possui prompts, dicas e mensagens colocados onde os olhos do usuário possivelmente estejam posicionados?**

a) Sim

b) Não

**2. O ambiente de visualização utiliza do espaço em branco para criar simetria e conduzir o olho do usuário na direção apropriada?**

a) Sim

b) Não

**3. O ambiente de visualização agrupa os itens em zonas lógicas, além de utilizar os cabeçalhos para distinguir tais zonas?**

a) Sim

b) Não

**4. O ambiente de visualização utiliza do realce de cores para chamar atenção do usuário?**

a) Sim

b) Não

**5. O ambiente de visualização possui um código de cores consistente em todo o sistema?**

a) Sim

b) Não

## **Questionário sobre flexibilidade e eficiência de uso**

**1. O ambiente de visualização possui um sistema de navegação simples e visível?**

a) Sim

b) Não

**2. O ambiente de visualização permite que usuários possam selecionar itens de uma lista curta (sete itens ou menos) movendo o cursor?**

a) Sim

b) Não

**3. O ambiente de visualização disponibiliza a opção de clicar em campos ou usar atalhos de teclado por meio de dispositivos apontadores?**

a) Sim

b) Não



**4. O ambiente de visualização oferece a opção de clicar diretamente ou usar atalhos de teclado nos itens de determinado menu?**

- a) Sim
- b) Não

## **Questionário sobre design estético e minimalista**

**1. O ambiente de visualização exibe em tela apenas as informações essenciais para a tomada de decisão?**

- a) Sim
- b) Não

**2. O ambiente de visualização utiliza de objetos grandes, fontes em negrito e áreas simples para distinguir as seções?**

- a) Sim
- b) Não

**3. O ambiente de visualização possui campos com rótulos breves, familiares e descritivos?**

- a) Sim
- b) Não

**4. O ambiente de visualização possui um layout visual bem desenhado?**

- a) Sim
- b) Não

**5. O ambiente de visualização evita elementos de dados desnecessários em cada tela?**

- a) Sim
- b) Não

## **Questionário sobre organização espacial**

**1. O ambiente de visualização disponibiliza seus elementos de forma que todos são claros e visíveis?**

- a) Sim
- b) Não

**2. O ambiente de visualização organiza as informações de forma lógica?**

- a) Sim
- b) Não

**3. O ambiente de visualização fornece detalhes sobre o contexto e seus elementos de dados?**

- a) Sim
- b) Não

## **Questionário sobre codificação de informação**

**1. O ambiente de visualização possui símbolos apropriados para os dados representados?**

- a) Sim
- b) Não

**2. O ambiente de visualização utiliza de características realistas para representar dados ou elementos de informação?**

- a) Sim
- b) Não

## **Questionário sobre orientação e ajuda**

**1. O ambiente de visualização oferece ao usuário a possibilidade de controlar o nível de detalhe que vê em uma representação?**

- a) Sim
- b) Não

**2. O ambiente de visualização oferece ao usuário a possibilidade de refazer/desfazer suas ações?**

- a) Sim
- b) Não

**3. O ambiente de visualização disponibiliza de recursos para que o usuário possa visualizar o caminho que seguiu para navegar até determinado ponto da aplicação?**

- a) Sim
- b) Não