



**UNIVERSIDADE FEDERAL DO CEARÁ  
CENTRO DE CIÊNCIAS  
DEPARTAMENTO DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO  
MESTRADO ACADÊMICO EM CIÊNCIA DA COMPUTAÇÃO**

**BRUNA NEVES DE OLIVEIRA**

**INVESTIGAÇÃO DA ADAPTAÇÃO DE TESTES DE USABILIDADE PARA AVALIAR  
FERRAMENTAS DE VISUALIZAÇÃO DE REDES MULTIVARIADAS**

**FORTALEZA  
2024**

BRUNA NEVES DE OLIVEIRA

INVESTIGAÇÃO DA ADAPTAÇÃO DE TESTES DE USABILIDADE PARA AVALIAR  
FERRAMENTAS DE VISUALIZAÇÃO DE REDES MULTIVARIADAS

Dissertação apresentada ao Curso de Mestrado Acadêmico em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação. Área de Concentração: Visualização de Dados.

Orientadora: Profa. Dra. Emanuele Marques Rodrigues Santos.

Coorientador: Prof. Dr. Fernando Antonio Mota Trinta.

FORTALEZA

2024

Dados Internacionais de Catalogação na Publicação  
Universidade Federal do Ceará  
Sistema de Bibliotecas

Gerada automaticamente pelo módulo Catalog, mediante os dados fornecidos pelo(a) autor(a)

---

O45i Oliveira, Bruna Neves de.

Investigação da adaptação de testes de usabilidade para avaliar ferramentas de visualização de redes multivariadas / Bruna Neves de Oliveira. – 2024.

122 f. : il. color.

Dissertação (mestrado) – Universidade Federal do Ceará, Centro de Ciências, Programa de Pós-Graduação em Ciência da Computação, Fortaleza, 2024.

Orientação: Prof. Dr. Emanuele Marques Rodrigues Santos.

Coorientação: Prof. Dr. Fernando Antonio Mota Trinta.

1. Avaliação com usuários. 2. Redes multivariadas. 3. Taxonomia de tarefas. I. Título.

CDD 005

---

**BRUNA NEVES DE OLIVEIRA**

**INVESTIGAÇÃO DA ADAPTAÇÃO DE TESTES DE USABILIDADE PARA AVALIAR  
FERRAMENTAS DE VISUALIZAÇÃO DE REDES MULTIVARIADAS**

Dissertação apresentada ao Curso de Mestrado Acadêmico em Ciência da Computação do Programa de Pós-Graduação em Ciência da Computação do Centro de Ciências da Universidade Federal do Ceará, como requisito parcial à obtenção do título de mestre em Ciência da Computação. Área de Concentração: Visualização de Dados.

Aprovada em: 30/08/2024.

**BANCA EXAMINADORA**

---

Profa. Dra. Emanuele Marques Rodrigues Santos (Orientadora)  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Fernando Antonio Mota Trinta (Coorientador)  
Universidade Federal do Ceará (UFC)

---

Profa. Dra. Ingrid Teixeira Monteiro  
Universidade Federal do Ceará (UFC)

---

Prof. Dr. Antonio José Melo Leite Júnior  
Universidade Federal do Ceará (UFC)

Aqueles que eu amo.

## **AGRADECIMENTOS**

Agradeço, inicialmente, ao meus pais e minhas irmãs, pois sem eles eu não teria chegado até aqui. Obrigada por tornarem este sonho uma realidade. Eu sou quem sou graças a vocês.

Agradeço à minha orientadora, Profa. Emanuele Santos, por ter aceitado ser minha orientadora e me guiado durante toda essa pesquisa. Manu, obrigada pelo seu excelente trabalho como professora e orientadora, você é uma mulher incrível.

A todos os participantes que contribuíram no experimento desta pesquisa, meus sinceros agradecimentos.

Agradeço também á Profa. Ingrid Monteiro, que me acompanhou durante praticamente todo o meu tempo na UFC: foi minha professora e primeira orientadora durante a graduação, e esteve presente em todas as bancas desta pesquisa. Obrigada, Ingrid, seus conhecimentos são valiosos demais.

Agradeço aos professores da banca revisora pela disponibilidade, pelas valiosas contribuições e pelas sugestões ao trabalho.

Sou muito grata aos meus amigos. Vocês são tantos que seria impossível mencionar cada um individualmente. Obrigada por sempre me oferecerem um ombro amigo e por estarem ao meu lado. Sei que vocês celebram este momento tanto quanto eu, e sou grata pela amizade de cada um de vocês.

Por fim, agradeço ao meu noivo, que me apoiou, comemorou comigo cada pequena vitória, escutou os meus lamentos e me abraçou nos momentos em que eu achava impossível terminar esse trabalho. Viver a vida com você me faz feliz demais.

"Believe." (Ted Lasso)

## RESUMO

As ferramentas de visualização de dados prometem aprimorar a tomada de decisão ao oferecer novas formas de compreender o mundo. Entretanto, fatores da experiência humana e a forma como os dados são apresentados influenciam significativamente a interação dos usuários com esses sistemas. A visualização de redes multivariadas, que envolve estruturas de dados abstratas contendo itens de dados, suas relações e propriedades associadas, é amplamente utilizada em diversas áreas. No entanto, as pesquisas sobre a avaliação de visualizações de redes multivariadas têm se concentrado majoritariamente em dois aspectos de usabilidade: eficiência e eficácia em tarefas relacionadas à topologia da rede, frequentemente utilizando redes fictícias e usuários sem especialização no domínio dos dados. Este trabalho apresenta uma investigação da adaptação de testes de usabilidade para incluir tarefas de análise de rede baseadas em taxonomia e domínio específico, com ênfase em participantes especialistas no domínio e dados reais. O objetivo é avaliar não apenas o desempenho da visualização, mas também a usabilidade da ferramenta de forma mais abrangente. O método proposto foi aplicado para avaliar uma ferramenta web na área de patologia computacional, que utiliza uma rede multivariada para representar lesões e doenças renais. Os resultados indicam que a avaliação foi eficaz na identificação de problemas de interação, na compreensão das expectativas dos usuários e na inclusão de novas funcionalidades na ferramenta. Além disso, são discutidas lições aprendidas ao longo do processo que são valiosas para melhorias na avaliação da usabilidade de ferramentas de visualização de dados em geral.

**Palavras-chave:** avaliação com usuários; redes multivariadas; taxonomia de tarefas.

## **ABSTRACT**

Data visualization tools promise to enhance decision-making by offering new ways to understand the world. However, factors related to human experience and the way data is presented significantly influence users' interaction with these systems. The visualization of multivariate networks, which involves abstract data structures containing data items, their relationships, and associated properties, is widely used across various fields. However, research on the evaluation of multivariate network visualizations has primarily focused on two usability aspects: efficiency and effectiveness in tasks related to network topology, often using fictitious networks and users without domain-specific expertise. This work presents an investigation into adapting usability testing to include network analysis tasks based on taxonomy and specific domains, with an emphasis on domain-expert participants and real data. The goal is to assess not only the visualization performance but also the usability of the tool in a more comprehensive manner. The proposed method was applied to evaluate a web tool in the field of computational pathology, which uses a multivariate network to represent renal lesions and diseases. The results indicate that the evaluation was effective in identifying interaction issues, understanding user expectations, and including new functionalities in the tool. Additionally, lessons learned throughout the process are discussed, which are valuable for improving the usability evaluation of data visualization tools in general.

**Keywords:** user evaluation; multivariate networks; task taxonomy.

## LISTA DE FIGURAS

Figura 1 – Exemplo de rede multivariada. Neste exemplo, a rede conecta artigos de acordo com sua semelhança fornecendo uma visão geral de um campo acadêmico. A intensidade da cor do nó representa o ano de publicação do artigo e o tamanho representa o número de citações dele. . . . .	16
Figura 2 – Topologia de uma rede. Representação de uma estrutura de conexões de nós	22
Figura 3 – Representação figurativa das abordagens das visualizações de redes multivariadas (as redes representadas em cada abordagem são diferentes) . . . . .	23
Figura 4 – Diagrama nó-link. Os dados baseiam-se na co-ocorrência de personagens em <i>Les Misérables</i> , de Victor Hugo, compilados por Donald Knuth. A rede possui 77 nós e 254 arestas. O tamanho do nó representa a quantidade de co-ocorrências e as cores representam os personagens que apareceram no mesmo capítulo. . . . .	24
Figura 5 – Matriz de Adjacência. Os dados da rede são os mesmos da Figura 4. As cores também representam os personagens que apareceram no mesmo capítulo e as células mais escuras indicam personagens que co-ocorreu com mais frequência. . . . .	24
Figura 6 – Estruturas topológicas que são consideradas como alvos para tarefas por Nobre <i>et al.</i> (2019). . . . .	27
Figura 7 – Resumo e relação entre as taxonomias de tarefas especializadas para visualização em redes apresentadas neste capítulo . . . . .	28
Figura 8 – Fluxo de execução da abordagem proposta para avaliação de ferramenta de visualização de rede . . . . .	47
Figura 9 – Comparação entre a abordagem proposta e as atividades do teste de usabilidade definidas por Barbosa <i>et al.</i> (2021) . . . . .	54
Figura 10 – Imagem de lâmina de biópsia exibindo algumas estruturas histológicas já segmentadas: dentre as quais glomérulos saudáveis, escleróticos e fibrose intersticial e atrofia tubular (IFTA) . . . . .	56
Figura 11 – Captura de tela do PS-iMAP mostrando seus três componentes visuais: uma rede multivariada de lesões e doenças representada em um diagrama nó-link, um painel para exibição de trabalhos relacionados de patologia computacional e uma galeria de imagens de lesões. . . . .	57

Figura 12 – Captura de tela do componente visual da rede multivariada. No componente há um campo de busca por lesão ou doença, uma barra de ferramentas e um diagrama nó-link representando a relação entre lesões e doenças renais . . .	58
Figura 13 – Captura de tela do PS-iMAP após a seleção do nó <i>SCLEROSIS, SEGMENTAL</i> mostrando o <i>card</i> de detalhes com uma imagem em miniatura e uma descrição da lesão selecionada. No componente da rede, o modo focado está ativado, destacando-se apenas o nó selecionado e os seus vizinhos imediatos. No painel para exibição de trabalhos relacionados e na galeria de imagens estão sendo exibidos somente os resultados referentes à lesão <i>SCLEROSIS, SEGMENTAL</i> . . . . .	60
Figura 14 – Captura de tela do Atlas. A imagem mostra as características da Nefropatia por IgA. . . . .	61
Figura 15 – Captura de tela do componente de trabalhos relacionados em patologia computacional . . . . .	61
Figura 16 – Captura de tela da galeria de lesões do PS-iMAP. . . . .	62
Figura 17 – Fluxo de execução da avaliação realizada no PS-iMAP. No planejamento da avaliação, os participantes foram separados em dois grupos: com e sem treinamento. . . . .	63
Figura 18 – Resumo do perfil dos participantes . . . . .	73
Figura 19 – Tempo de execução das tarefas bem-sucedidas . . . . .	83

## LISTA DE TABELAS

Tabela 1 – Exemplos de tarefas de análise de redes multivariadas compostas por tarefas básicas e a estrutura topológica . . . . .	27
Tabela 2 – Atividades do teste de usabilidade . . . . .	35
Tabela 3 – Lista de tarefas e instruções dadas aos participantes . . . . .	44
Tabela 4 – Relação das características das avaliações . . . . .	45
Tabela 5 – Perguntas que a avaliação buscou responder . . . . .	64
Tabela 6 – Lista de instrução da tarefa e o seu tipo: análise de rede ou interação com a ferramenta . . . . .	66
Tabela 7 – Lista das tarefas de análise de rede, estrutura topológica e tipo de tarefa de acordo com a taxonomia de Nobre <i>et al.</i> (2019) . . . . .	67
Tabela 8 – Tempo da duração da tarefa de exploração livre, a quantidade total de comentários e a categoria do primeiro comentário realizado durante a tarefa por participante. . . . .	74
Tabela 9 – Primeira ação ou clique que o participante realizou ao interagir com o PS-iMAP	75
Tabela 10 – Tempo da conclusão em minutos e o grau de sucesso da execução das tarefas. As tarefas executadas com sucesso são representadas em verde e as não sucedidas em vermelho . . . . .	76
Tabela 11 – Resumo das sugestões de melhoria que os participantes apontaram para a ferramenta . . . . .	81
Tabela 12 – Resumo das sugestões de novas funcionalidades . . . . .	81

## LISTA DE ABREVIATURAS E SIGLAS

AgT	<i>Attributes for a given Topological structure</i>
CEP	Comitê de Ética em Pesquisa
ENIAC	<i>Electronic Numerical Integrator and Calculator</i>
GUI	<i>Graphical User Interface</i>
IHC	Interação Humano-Computador
IHD	Interação Humano-Dados
ISO	<i>International Organization for Standardization</i>
MILCs	<i>Multi-dimensional In-depth Long-term Case studies</i>
MTurk	Amazon Mechanical Turk
SEQ	<i>Single Ease Question</i>
SGB	<i>Stanford Graph Base</i>
SUS	<i>System Usability Scale</i>
TCLE	Termo de Consentimento Livre e Esclarecido
TgA	<i>Topology for given Attributes</i>
USE	<i>Usefulness, Satisfaction, and Ease of use</i>
UX	<i>User Experience</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>15</b>
<b>1.1</b>	<b>Objetivos . . . . .</b>	<b>17</b>
<b>1.2</b>	<b>Etapas de Pesquisa . . . . .</b>	<b>18</b>
<b>1.3</b>	<b>Organização do Trabalho . . . . .</b>	<b>18</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>19</b>
<b>2.1</b>	<b>Visualização de Dados . . . . .</b>	<b>19</b>
<b>2.2</b>	<b>Redes Multivariadas . . . . .</b>	<b>21</b>
<b>2.2.1</b>	<b><i>Características de Redes Multivariadas</i> . . . . .</b>	<b>22</b>
<b>2.2.2</b>	<b><i>Visualização de Redes Multivariadas</i> . . . . .</b>	<b>22</b>
<b>2.2.3</b>	<b><i>Tarefas de Análise</i> . . . . .</b>	<b>23</b>
<b>2.3</b>	<b>Avaliação de Interação Humano-Computador . . . . .</b>	<b>28</b>
<b>2.3.1</b>	<b><i>Usabilidade</i> . . . . .</b>	<b>31</b>
<b>2.3.2</b>	<b><i>Teste de usabilidade</i> . . . . .</b>	<b>33</b>
<b>2.4</b>	<b>Considerações Finais . . . . .</b>	<b>36</b>
<b>3</b>	<b>TRABALHOS RELACIONADOS . . . . .</b>	<b>38</b>
<b>3.1</b>	<b>Métodos de avaliação de visualização de dados . . . . .</b>	<b>38</b>
<b>3.2</b>	<b>Avaliação de visualização de redes multivariadas . . . . .</b>	<b>41</b>
<b>3.3</b>	<b>Considerações Finais . . . . .</b>	<b>46</b>
<b>4</b>	<b>ADAPTAÇÃO DE TESTES DE USABILIDADE . . . . .</b>	<b>47</b>
<b>4.1</b>	<b>Definição dos objetivos da avaliação . . . . .</b>	<b>47</b>
<b>4.2</b>	<b>Escolha da modalidade de avaliação . . . . .</b>	<b>48</b>
<b>4.3</b>	<b>Planejamento . . . . .</b>	<b>49</b>
<b>4.3.1</b>	<b><i>Identificar as tarefas de análise da rede</i> . . . . .</b>	<b>49</b>
<b>4.3.2</b>	<b><i>Identificar as tarefas de interação</i> . . . . .</b>	<b>50</b>
<b>4.3.3</b>	<b><i>Definir o perfil e a quantidade de participantes</i> . . . . .</b>	<b>50</b>
<b>4.3.4</b>	<b><i>Preparar os materiais de apoio</i> . . . . .</b>	<b>51</b>
<b>4.3.5</b>	<b><i>Teste-piloto</i> . . . . .</b>	<b>51</b>
<b>4.3.6</b>	<b><i>Recrutar os participantes</i> . . . . .</b>	<b>52</b>
<b>4.4</b>	<b>Condução da avaliação . . . . .</b>	<b>52</b>
<b>4.5</b>	<b>Análise dos dados . . . . .</b>	<b>53</b>

<b>4.6</b>	<b>Considerações Finais . . . . .</b>	54
<b>5</b>	<b>PS-IMAP . . . . .</b>	55
<b>5.1</b>	<b>Rede multivariada de lesões e doenças . . . . .</b>	56
<b>5.2</b>	<b>Painel de trabalhos relacionados . . . . .</b>	59
<b>5.3</b>	<b>Galeria de imagens de lesões . . . . .</b>	59
<b>5.4</b>	<b>Considerações Finais . . . . .</b>	59
<b>6</b>	<b>APLICAÇÃO DO TESTE DE USABILIDADE: ESTUDO DE CASO</b>	
	<b>COM O PS-IMAP . . . . .</b>	63
<b>6.1</b>	<b>Avaliação do PS-iMAP . . . . .</b>	63
<b>6.1.1</b>	<b><i>Definição dos objetivos da avaliação . . . . .</i></b>	64
<b>6.1.2</b>	<b><i>Escolha da modalidade da avaliação . . . . .</i></b>	65
<b>6.1.3</b>	<b><i>Planejamento . . . . .</i></b>	65
<b>6.1.3.1</b>	<b><i>Descrição das tarefas . . . . .</i></b>	65
<b>6.1.3.1.1</b>	<b>Tarefas de análise da rede . . . . .</b>	67
<b>6.1.3.1.2</b>	<b>Tarefas de interação . . . . .</b>	67
<b>6.1.3.2</b>	<b><i>Definição do perfil e da quantidade de participantes . . . . .</i></b>	68
<b>6.1.3.3</b>	<b><i>Preparação dos materiais de apoio . . . . .</i></b>	68
<b>6.1.3.4</b>	<b><i>Testes-piloto . . . . .</i></b>	69
<b>6.1.3.5</b>	<b><i>Recrutamento dos participantes . . . . .</i></b>	70
<b>6.1.4</b>	<b><i>Condução da avaliação . . . . .</i></b>	70
<b>6.1.5</b>	<b><i>Análise dos dados . . . . .</i></b>	72
<b>6.2</b>	<b>Análise dos Resultados da Avaliação . . . . .</b>	72
<b>6.2.1</b>	<b><i>Entrevista estruturada pré-teste . . . . .</i></b>	72
<b>6.2.2</b>	<b><i>Exploração livre . . . . .</i></b>	74
<b>6.2.3</b>	<b><i>Observação de uso . . . . .</i></b>	75
<b>6.2.4</b>	<b><i>Entrevista semiestruturada pós-teste . . . . .</i></b>	79
<b>6.2.5</b>	<b><i>Resultados Quantitativos . . . . .</i></b>	82
<b>6.2.6</b>	<b><i>Discussão . . . . .</i></b>	82
<b>6.3</b>	<b>Considerações Finais . . . . .</b>	84
<b>7</b>	<b>LIÇÕES APRENDIDAS . . . . .</b>	86
<b>7.1</b>	<b>Métodos de avaliação de visualização de dados . . . . .</b>	86
<b>7.2</b>	<b>Condições que refletem a realidade . . . . .</b>	87

<b>7.3</b>	<b>Tarefas de análise de visualização de dados . . . . .</b>	<b>87</b>
<b>7.4</b>	<b>Teste de usabilidade . . . . .</b>	<b>88</b>
<b>7.5</b>	<b>Considerações finais . . . . .</b>	<b>89</b>
<b>8</b>	<b>CONCLUSÃO E TRABALHOS FUTUROS . . . . .</b>	<b>90</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>92</b>
	<b>APÊNDICE A –TERMOS DE CONSENTIMENTO LIVRE E ESCLARECIDO . . . . .</b>	<b>98</b>
	<b>APÊNDICE B –ENTREVISTA ESTRUTURADA PRÉ-TESTE . . . . .</b>	<b>102</b>
	<b>APÊNDICE C –CENÁRIO COM TAREFAS PARA PATOLOGISTAS . . . . .</b>	<b>103</b>
	<b>APÊNDICE D –CENÁRIO COM TAREFAS PARA CIENTISTA DA COMPUTAÇÃO . . . . .</b>	<b>105</b>
	<b>APÊNDICE E –ENTREVISTA SEMIESTRUTURADA PÓS-TESTE . . . . .</b>	<b>107</b>
	<b>APÊNDICE F –ROTEIRO . . . . .</b>	<b>108</b>
	<b>APÊNDICE G –ANÁLISE DE CONTEÚDO . . . . .</b>	<b>110</b>
	<b>ANEXO A –PARECER DO COMITÊ DE ÉTICA EM PESQUISA . . . . .</b>	<b>113</b>

## 1 INTRODUÇÃO

A tecnologia está em constante mudança e evolução, buscando continuamente novas formas de aprimorar a interação entre as pessoas, o ambiente e até mesmo os dados. A era moderna é caracterizada pela promessa de uma melhor tomada de decisão por meio do acesso a mais dados do que jamais se tinha visto (MUNZNER, 2014). A disponibilidade dessas enormes quantidades de dados, juntamente com as ferramentas estatísticas para “mastigar” esses números, oferece toda uma nova maneira de compreender o mundo (SANTAELLA; KAUFMAN, 2021).

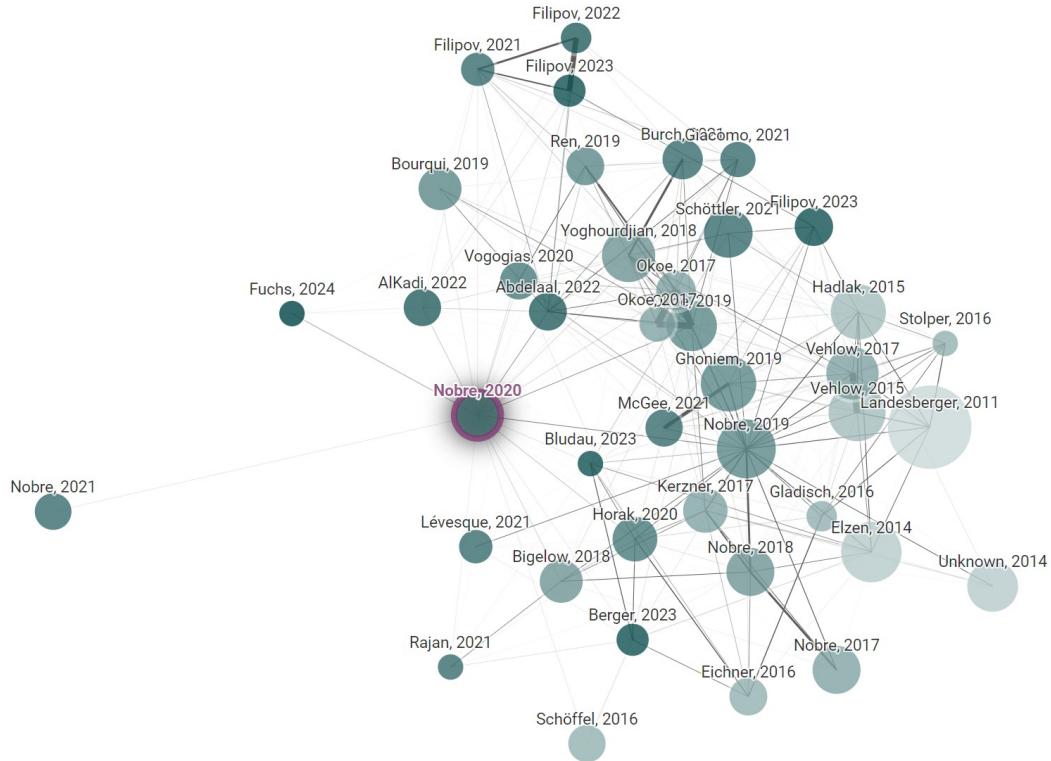
Gráficos, diagramas e ilustrações só se tornaram amplamente disponíveis como ferramentas de pensamento visual nos últimos duzentos anos (WARE, 2010). Atualmente, essas ferramentas são amplamente utilizadas em diversas áreas, incluindo negócios, ciência, análise de dados, jornalismo, e muitas outras, sendo essenciais para transformar informações complexas em representações gráficas claras e comprehensíveis, ajudando a identificar padrões, tendências e *insights*, facilitando a análise, interpretação, tomada de decisão e a comunicação dos resultados.

Um tipo de visualização de dados são as redes multivariadas, que envolvem uma estrutura de dados abstrata e onipresente, definida, em sua forma mais simples, como a combinação de um conjunto de pontos de dados (nós) e as relações entre eles (arestas) (FILIPOV *et al.*, 2023). Uma rede multivariada possui atributos associados a nós e arestas (KERREN *et al.*, 2014) e são aplicadas em várias disciplinas, incluindo: redes sociais<sup>1</sup>, redes que modelam processos celulares na biologia, redes que fornecem uma visão geral de um campo acadêmico (exemplo na Figura 1), além de árvores que descrevem relações evolutivas entre espécies.

Estudar cenários de visualização é difícil porque fatores da experiência humana e habilidade de análise, bem como os dados e sua apresentação influenciam o desempenho dentro de um sistema de visualização (ISENBERG *et al.*, 2008). Por isso, há a necessidade de que sejam realizadas avaliações com usuários. Os relatórios de estudos de usabilidade e experimentos controlados são úteis para entender o potencial e as limitações das ferramentas de visualização (PLAISANT, 2004), garantindo que elas atendam às necessidades dos usuários e forneçam informações relevantes de forma clara e acessível. No entanto, a pesquisa existente em estudos de avaliação de visualização de rede, até agora, tem focado majoritariamente em dois aspectos de usabilidade: eficiência e eficácia em tarefas relacionadas à topologia da rede, frequentemente utilizando redes fictícias ou geradas aleatoriamente e usuários sem conhecimento no domínio

<sup>1</sup> Rede social refere-se a uma representação de como diferentes elementos (geralmente pessoas, organizações, ou entidades) estão conectados ou interagem entre si.

Figura 1 – Exemplo de rede multivariada. Neste exemplo, a rede conecta artigos de acordo com sua semelhança fornecendo uma visão geral de um campo acadêmico. A intensidade da cor do nó representa o ano de publicação do artigo e o tamanho representa o número de citações dele.



Fonte: A Autora a partir de Connected Papers (2024).

dos dados.

Para melhorar a avaliação de sistemas de visualização, além dos usuários, é importante que o sistema seja avaliado com *datasets* reais e tarefas realistas (PLAISANT, 2004). As taxonomias de tarefas organizam diferentes técnicas e métodos utilizados para estudar e compreender visualizações de dados, traduzindo as tarefas dos usuários de um contexto específico de domínio para uma forma abstrata. Isso permite identificar semelhanças e diferenças entre as necessidades dos usuários em diversos contextos de uso no mundo real. Existem várias taxonomias de tarefas de visualização independentes do tipo de dados, inclusive taxonomias especializadas para redes. Entretanto, as avaliações com tarefas normalmente focam exclusivamente no aspecto de desempenho do sistema, resumindo-se à coleta de métricas de taxas de erros e tempo de execução para realizar as tarefas (SAKET *et al.*, 2016). Na literatura de visualização de redes, por exemplo, estudos concentram-se na comparação de representações visuais de redes e poucos utilizam alguma taxonomia de tarefas (JIANU *et al.*, 2014; REN *et al.*, 2019; NOBRE *et al.*, 2020). Muitos estudos que compararam os tipos de visualização de redes geralmente se concentram na eficiência dos algoritmos de layout ou nas diferenças estéticas (REN *et al.*, 2019),

e não na experiência do usuário.

É preciso explorar novos métodos de avaliação que vão além dessas métricas de desempenho estabelecidas. Na área de Interação Humano-Computador (IHC), a avaliação é necessária para a certificação de que os usuários podem vir a utilizar o produto e apreciá-lo (PREECE *et al.*, 2005), buscando entender como os usuários interagem com a tecnologia e identificar possíveis problemas ou melhorias para tornar a experiência mais satisfatória e produtiva, com alta qualidade de uso.

## 1.1 Objetivos

Inicialmente a pesquisa buscou entender como é avaliada a usabilidade da visualização de dados e de redes multivariadas. Foram então realizadas pesquisas teóricas com o objetivo de levantar e discutir as principais características das avaliações. No caminho da pesquisa, além de alcançar esse entendimento, surgiu a seguinte pergunta: “Como adaptar o teste de usabilidade para avaliar ferramentas com visualização de redes multivariadas de forma mais abrangente?”

Para responder à pergunta acima, propõem-se os seguintes objetivos específicos:

1. Fazer o levantamento de metodologias de avaliação de visualização de dados e de IHC, de redes multivariadas e de taxonomias de tarefas de análise;
2. Investigar uma abordagem de avaliação que possa ser utilizada por designers e desenvolvedores de ferramentas de visualização de redes multivariadas;
3. Analisar e discutir os resultados da abordagem.

Dessa forma, este trabalho apresenta uma investigação da adaptação do método de IHC teste de usabilidade para incluir tarefas de análise de rede baseadas em taxonomia e domínio específico, com ênfase em participantes especialistas no domínio e dados reais. A intenção é avaliar não apenas o desempenho da visualização, mas também a usabilidade da ferramenta de forma mais abrangente em condições que refletem a realidade. A abordagem proposta foi aplicada para avaliar o PS-iMAP, uma ferramenta web que utiliza uma rede multivariada com relações entre lesões e doenças renais para melhorar a comunicação e a colaboração entre patologistas e pesquisadores de ciência da computação em projetos na área de patologia computacional.

Espera-se também que este trabalho ilustre a importância do uso de métodos de IHC para profissionais da área de visualização de dados.

## 1.2 Etapas de Pesquisa

Inicialmente para a execução desse trabalho, foram realizadas pesquisas teóricas para o levantamento das técnicas de avaliação de visualização de dados, de IHC e de redes multivariadas. Também foram levantadas as taxonomias de tarefas de análise de redes, etapa essencial para a definição das tarefas de análise de rede que devem ser utilizadas durante a avaliação. Em seguida, foi elaborada uma abordagem metodológica com adaptações do teste de usabilidade para avaliar a usabilidade tanto da ferramenta como da visualização, incorporando as tarefas de análise de rede baseadas em taxonomia. Ao final, a metodologia foi aplicada para avaliar o PS-iMAP. Dessa avaliação foi possível extrair conclusões sobre todo o processo.

## 1.3 Organização do Trabalho

O restante deste documento está organizado da seguinte forma: o Capítulo 2 apresenta os conceitos teóricos necessários para a compreensão deste trabalho. No Capítulo 3, são discutidos os trabalhos relacionados. No capítulo 4, é detalhada a adaptação de testes de usabilidade para a avaliação de ferramentas de visualização de redes multivariadas. No Capítulo 5, descreve-se a aplicação PS-iMAP. O Capítulo 6 mostra como a adaptação proposta foi aplicada e os resultados da avaliação do PS-iMAP. No Capítulo 7, discutem-se as lições aprendidas e, por fim, o Capítulo 8 traz as conclusões e as perspectivas de trabalhos futuros.

## 2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são mostrados os conceitos mais importantes para a compreensão do trabalho: Visualização de dados, com a intenção de compreender os princípios e práticas envolvidas na criação, interpretação e comunicação de informações; Redes Multivariadas, com a finalidade de conhecer as principais abordagens e as taxonomias de tarefas de análise; e Avaliação de IHC, para o entendimento dos métodos de avaliação.

### 2.1 Visualização de Dados

As primeiras visualizações surgiram em diagramas geométricos, em tabelas de posições de estrelas e outros corpos celestes e na confecção de mapas para auxiliar na navegação e exploração<sup>2</sup> (CHEN *et al.*, 2007, p. 18). Diagramas, mapas, páginas da web, gráficos informativos, instruções visuais e ilustrações técnicas nos ajudam a resolver problemas por meio de um processo de pensamento visual (WARE, 2010).

Uma apresentação visual elementar consiste em um conjunto de marcas (como Pontos, Linhas, Áreas, Superfícies ou Volumes), seus canais ou propriedades (como Cor e Tamanho) e sua posição no espaço e no tempo (como o plano X, Y em computação gráfica clássica, XYZT ou 3D mais tempo (T) em visualização da informação) (CARD; MACKINLAY, 1997). Mesmo codificações visuais complexas podem ser divididas em componentes que podem ser analisados em termos da estrutura de suas marcas e canais (MUNZNER, 2014). A maioria das análises usando gráficos faz uso de um conjunto padrão de ferramentas gráficas, por exemplo, gráficos de dispersão, gráficos de barras e histogramas (CHEN *et al.*, 2007, p. 18).

Uma variedade de diferentes tipos de dados abstratos é abordada na pesquisa de visualização de dados (por exemplo, numérico, ordinal, categórico), que podem ser organizados de diferentes maneiras: por exemplo, em forma linear, tabular ou rede (KERREN *et al.*, 2014). Representações comuns de dados estatísticos como os gráficos de pizza, gráficos de barras, ou gráficos de dispersão, são visualizações de dados numéricos abstratos.

Os sistemas de visualização baseados em computador fornecem representações visuais de conjuntos de dados projetados para ajudar as pessoas a realizar tarefas com mais eficiência (MUNZNER, 2014). Instrumentos convencionalmente custosos e lentos (questionários, enquetes, cálculos estatísticos) dão lugar a ferramentas mais ágeis e simples (sistemas automatizados de

<sup>2</sup> A ideia de coordenadas foi usada por antigos egípcios por pelo menos 200 A.C. e a projeção de mapa de uma Terra esférica serviria como padrão de referência até o século XIV (CHEN *et al.*, 2007)

coleta, registro e visualização) (BRUNO, 2012).

A visualização de dados é uma parte importante de qualquer análise estatística e serve para muitos propósitos diferentes. A visualização é útil para entender a estrutura geral e a natureza dos dados, como os tipos de variáveis contidas nos dados (categóricas, numéricas, textuais etc.), suas faixas de valores e o equilíbrio entre elas. A visualização é útil para detectar dados ausentes e também pode ajudar a identificar observações extremas e discrepantes. Além disso, tendências e padrões desconhecidos nos dados são frequentemente descobertos com a ajuda da visualização. (CHEN *et al.*, 2007, p. 874).

As técnicas de visualização são de grande importância em uma ampla gama de áreas de aplicação. O volume, velocidade, heterogeneidade e a complexidade dos dados disponíveis tornam extremamente difícil para os humanos explorar e analisar os dados (BIKAKIS, 2018).

A visualização de dados muda a maneira de como as pessoas experimentam as informações e a maneira de como vivem (APARICIO; COSTA, 2015). A maioria dos projetos de visualização é para auxiliar na tomada de decisões e servir como ferramentas que aumentam a cognição (SADIKU *et al.*, 2016). Ferramentas de visualização de dados são utilizadas em diversas áreas como negócios, ciência, análise de dados, jornalismo entre outras, sendo essenciais para transformar informações complexas em representações gráficas claras e comprehensíveis, ajudando a identificar padrões, tendências e *insights*, facilitando a análise, interpretação, tomada de decisão e a comunicação dos resultados.

Os dispositivos que o usuário usa, as redes por quais eles se conectam e as interações que usam essas tecnologias geram rastros consideráveis de dados. Esses dados pessoais são acumulados por muitas organizações diferentes e eles são processados usando algoritmos cada vez mais sofisticados para medir e inferir características da vida do usuário. Essa prática crescente de acumulação de dados e a importância crescente desses dados e inferências extraídas deles para o nosso cotidiano impulsionam a necessidade do estudo de Interação Humano-Dados (IHD) (MORTIER *et al.*, 2014).

IHD é a área que aborda a manipulação humana, análise e *sensemaking*<sup>3</sup> de grandes conjuntos de dados não estruturados e complexos, entendendo seu significado, bem como considerando as partes interessadas e as fases do ciclo de vida dos dados (VICTORELLI *et al.*, 2020). A IHD coloca o humano no centro dos fluxos de dados, e está preocupada em fornecer mecanismos para interação explicitamente com esses sistemas e dados (HADDADI, 2016), estudando fenômenos relacionados à interação de pessoas com aplicações de software que manipulam dados pessoais e seu consequente impacto (HADDADI *et al.*, 2015).

---

<sup>3</sup> O termo "*sensemaking*" refere-se ao processo de dar sentido a situações complexas ou confusas.

No entanto, ainda há carência de estudos que proporcionem uma compreensão aprofundada da área (VICTORELLI *et al.*, 2020), por ser um conceito novo e os pesquisadores apresentam visões específicas do assunto. A IHD diz respeito à interação geralmente entre humanos, conjuntos de dados e análises, mas não ao estudo geral da interação com sistemas computacionais que é a IHC (MORTIER *et al.*, 2013).

Dentro dos conceitos apresentados até aqui, este trabalho foca na investigação da avaliação da usabilidade, utilizando conceitos de IHC, de ferramentas que possuem visualização de redes multivariadas, conceito abordado a seguir.

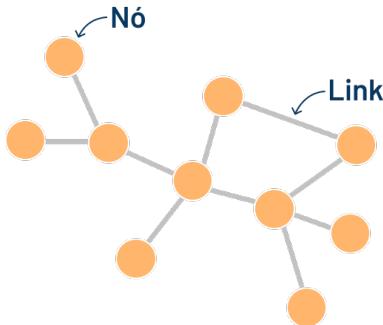
## 2.2 Redes Multivariadas

Uma rede é uma estrutura de dados abstrata e onipresente, definida, em sua forma mais simples, como a combinação de um conjunto de pontos de dados e as relações entre eles (FILIPOV *et al.*, 2023). As redes são utilizadas em diversas áreas e a sua utilização varia entre redes físicas, como redes elétricas, redes sociais, redes que modelam processos celulares na biologia, e árvores que descrevem relações evolutivas das espécies.

Uma rede multivariada é definida tendo duas importantes características: primeiro, os nós são conectados uns aos outros por meio de links (ou arestas), ou seja, existe uma estrutura topológica. Segundo, sendo multivariada, os nós e links têm atributos associados a eles, e esses atributos possuem um valor (PRETORIUS *et al.*, 2014). Uma rede multivariada é um tipo de dado abstrato que fornece desafios particulares para a comunidade de pesquisa em visualização de dados, como o desafio da escalabilidade - o que fazer quando as redes são tão grandes que não podem ser exibidas de forma eficaz (KERREN *et al.*, 2014). Ela permite a representação de dados relacionais complexos (armazenados na forma de uma rede) bem como a associação de atributos a esses dados (KERREN *et al.*, 2014). O objetivo fundamental com os dados de rede é entender a estrutura dessas interconexões; ou seja, a topologia da rede (MUNZNER, 2014) (Figura 2).

A maioria das redes da vida real é multivariada (NOBRE *et al.*, 2019). Seja o exemplo de uma rede social: os nós são pessoas, os atributos dos nós podem ser idade, nome e endereço, e os atributos de links podem ser o tipo de relação, a duração da relação e a frequência das interações. As redes multivariadas também são usadas para representar conhecimento de domínio abstrato, por exemplo, em engenharia de software, biologia, evolução, ciências ambientais, meteorologia e sociologia. Ao se visualizar topologia e atributos numa mesma visão,

Figura 2 – Topologia de uma rede. Representação de uma estrutura de conexões de nós



Fonte: Elaborado pela autora.

escolher codificações visuais<sup>4</sup> eficientes para um dos aspectos frequentemente interfere com a habilidade de visualizar o outro de forma efetiva (NOBRE *et al.*, 2019; YOGHOURDJIAN *et al.*, 2021), pois o número de canais visuais disponíveis, tais como cor, tamanho e posição, é limitado. Portanto, o uso de técnicas computacionais para ajudar na visualização efetiva de redes é uma área de estudo com ampla aplicabilidade, bem como potencial para ser altamente útil em outras áreas de pesquisa (KERREN *et al.*, 2014).

### **2.2.1 Características de Redes Multivariadas**

Em redes multivariadas, normalmente os nós são de um mesmo tipo (pessoas em redes sociais, por exemplo) e os atributos podem ser de tipos de dados variados como textuais, quantitativos, ordinais, nominais entre outros. No entanto, quando os nós e arestas podem ser de diferentes tipos (autores e documentos em redes de citações, por exemplo), a rede é dita heterogênea.

Em termos de tamanho das redes, são consideradas como pequenas, redes com menos de 100 nós, médias com 100 a 1.000 nós, e grandes com mais de 1.000 nós. Para atributos de nós, distingue-se entre poucos (menos de cinco) e muitos (mais de cinco). Já para as arestas, considera-se poucos com menos de três e muitos com três ou mais atributos (NOBRE *et al.*, 2019).

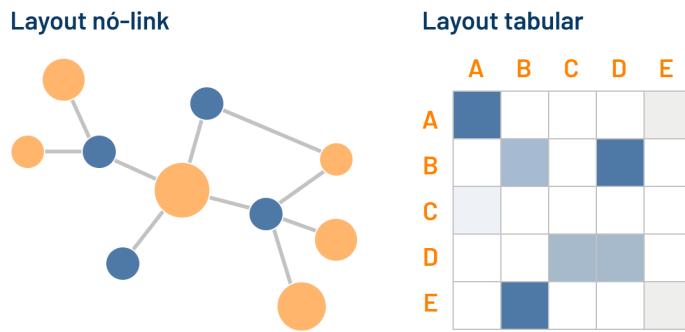
### **2.2.2 Visualização de Redes Multivariadas**

As abordagens de visualização de redes multivariadas são: layouts de nó-link e layouts tabulares (Figura 3). Os diagramas de nó-link são os gráficos mais comuns para repre-

<sup>4</sup> O núcleo do espaço de design das codificações visuais pode ser descrito como uma combinação de dois aspectos: as marcas que são os elementos gráficos e os canais visuais para controlar sua aparência (MUNZNER, 2014).

sentação de redes, os nós são desenhados com um ponto e os links como uma linha ou curva que conecta os nós (Figura 3, à esquerda). Nos layouts tabulares, os nós e/ou links são representados por colunas e/ou linhas de uma tabela, a abordagem mais conhecida são as matrizes de adjacência (Figura 3, à direita).

Figura 3 – Representação figurativa das abordagens das visualizações de redes multivariadas (as redes representadas em cada abordagem são diferentes)



Fonte: Elaborado pela autora.

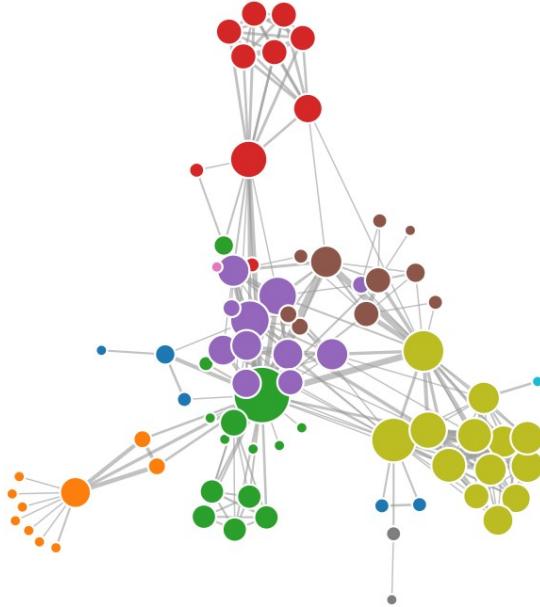
As Figuras 4 e 5 foram criadas a partir do mesmo conjunto de dados com as co-ocorrências dos personagens do livro *Les Misérables*, de Victor Hugo. O *dataset* foi criado por Donald Knuth em 1993, como parte da *Stanford Graph Base* (SGB) (ROSSI; AHMED, 2015). O *dataset* contém 77 nós correspondentes aos personagens do romance e 254 arestas representando os encontros entre esses personagens. Os personagens são coloridos conforme a co-ocorrência no mesmo capítulo.

Cada abordagem de visualização de redes multivariadas possui suas vantagens e desvantagens. A grande vantagem do diagrama nó-link é que, para redes pequenas, ele é bastante intuitivo e dá suporte a muitas tarefas de análise de redes multivariadas (explicadas na seção a seguir), especialmente tarefas relacionadas à topologia da rede. A sua principal desvantagem é que a partir de um certo número de nós e arestas, torna-se impossível uma visualização efetiva. Layouts tabulares são mais apropriados para visualização de redes grandes e densas, mas dificultam o entendimento da topologia da rede.

### 2.2.3 Tarefas de Análise

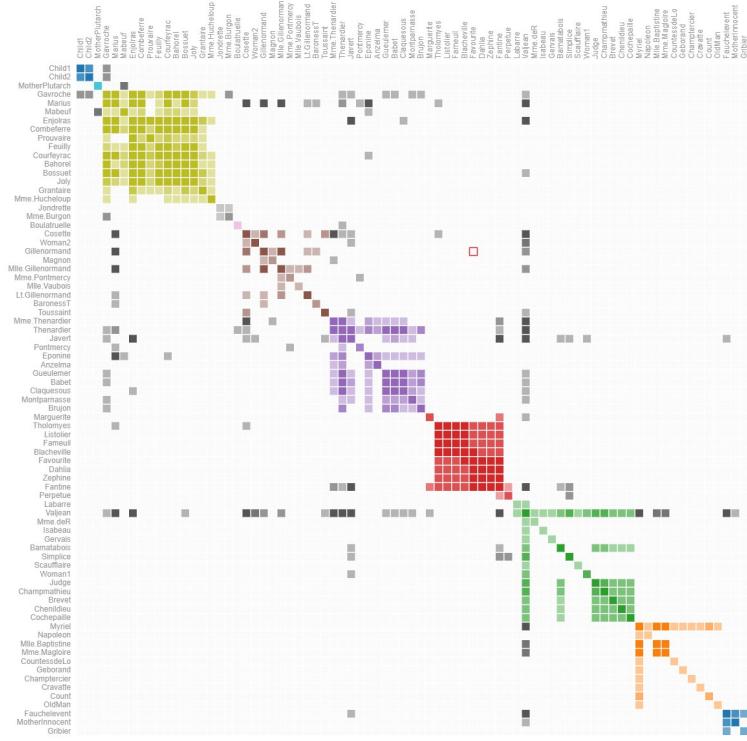
No contexto das redes multivariadas, o termo tarefa é associado à atividade que um usuário deseja realizar interagindo com uma representação visual de uma rede multivariada. Isso implica que há a intenção do usuário e que a rede foi apresentada visualmente, essa

Figura 4 – Diagrama nó-link. Os dados baseiam-se na co-ocorrência de personagens em *Les Misérables*, de Victor Hugo, compilados por Donald Knuth. A rede possui 77 nós e 254 arestas. O tamanho do nó representa a quantidade de co-ocorrências e as cores representam os personagens que apareceram no mesmo capítulo.



Fonte: Elaborado pela autora.

Figura 5 – Matriz de Adjacência. Os dados da rede são os mesmos da Figura 4. As cores também representam os personagens que apareceram no mesmo capítulo e as células mais escuras indicam personagens que co-ocorreram com mais frequência.



Fonte: Elaborado pela autora.

intenção é geralmente descrita como o objetivo de obter informações sobre os dados apresentados (PRETORIUS *et al.*, 2014). No caso de uma rede social, alguns exemplos de tarefas podem ser: “Encontre a pessoa com mais amigos”, “Encontre a amizade mais antiga de Maria”, “Selecione a pessoa com quem Maria mais interage”, entre outros.

Existem várias taxonomias de tarefas de visualização independentes de tipo de dados. Wehrend e Lewis (1990) propõem uma classificação dos métodos de visualização, considerando as entidades estudadas e as tarefas executadas nessas entidades. Eles listam 11 tarefas encontradas mais frequentemente: identificar, localizar, distinguir, categorizar, agrupar (*cluster*)<sup>5</sup>, (analisar) distribuição, classificar, comparar, (analisar) dentro e entre relações, associar, e correlacionar. Já Amar e Stasko (2005) apontam que as tarefas devem ser tarefas típicas que pessoas reais querem realizar usando as ferramentas de visualização, as tarefas podem incluir: tomar decisões complexas, aprender um domínio, identificar a natureza das tendências, prever o futuro, identificar os parâmetros do domínio, descobrir modelos correlatos, formular e verificar hipóteses, identificar o efeito das incertezas dos dados e identificar fontes de causalidade.

Também existem taxonomias de tarefas especializadas para visualização em redes. Uma tarefa de visualização de rede envolve a realização de uma atividade analítica em uma combinação de entidade (a “coisa” que está sendo estudada) e uma propriedade dessa entidade (PRETORIUS *et al.*, 2014). Nas redes multivariadas, as entidades são: nós, links, caminhos ou sequência de nós e links, e redes. As propriedades associadas são: topologia e atributos associados a nós e links. Voltando ao exemplo de uma rede social, ao analisar caminhos entre os nós, é possível obter conhecimento, mesmo que duas pessoas não tenham relação direta, se ambas possuírem uma relação com uma terceira pessoa em comum, existe uma relação indireta entre elas que pode (ou não) ser de interesse (PRETORIUS *et al.*, 2014).

Valiati *et al.* (2006) abordam três classes de tarefas: operacionais (relativas aos meios pelos quais a rede é apresentada e explorada), analíticas (relativas ao meio pelo qual a informação é extraída da rede) e cognitivas (relativas a compreensão de toda a rede), compreendidas em sete tarefas: identificar, determinar, visualizar, comparar, inferir, configurar e localizar. Lee *et al.* (2006) propõem quatro categorias de tarefas: baseadas em topologia, baseadas em atributos, navegação e visão geral, descrevendo tarefas comumente encontradas ao analisar redes e propondo uma série de tarefas para cada uma das categorias.

Baseado nas categorias e tarefas constituintes definidas por Valiati *et al.* (2006),

---

<sup>5</sup> Um *cluster* é um agrupamento baseado na similaridade de atributos, onde os itens dentro de um *cluster* são mais semelhantes entre si do que aqueles em outro *cluster* (MUNZNER, 2014).

Pretorius *et al.* (2014) classificam sua taxonomia da seguinte forma:

- Operacionais: visualizar, configurar;
- Analíticas: identificar, determinar, realocar, comparar; e
- Cognitiva: inferir.

Além disso, Pretorius *et al.* (2014) generalizaram a taxonomia de Lee *et al.* (2006) para atender outras entidades, como links e *clusters*. A partir disso, as tarefas também são categorizadas em: baseada em estrutura (topologia): adjacência, acessibilidade, conexão e conectividade; baseada em atributos: nós e arestas; navegação: seguir caminho e revisitar; e estimativa: entendimento e comparação. Pretorius *et al.* (2014) indicam que o entendimento do problema de domínio deve ser traduzido em tarefas de usuário e as tarefas de usuário devem ter relação direta com a visualização projetada, além disso, a compreensão das tarefas permite a avaliação da adequação da solução para resolver o problema.

Nobre *et al.* (2019) apresentam uma taxonomia de tarefas com foco específico em análise de rede multivariada, em que qualquer tarefa pode ser expressa como uma combinação de duas tarefas fundamentais, aplicadas a diferentes estruturas topológicas de uma rede. As duas tarefas básicas são:

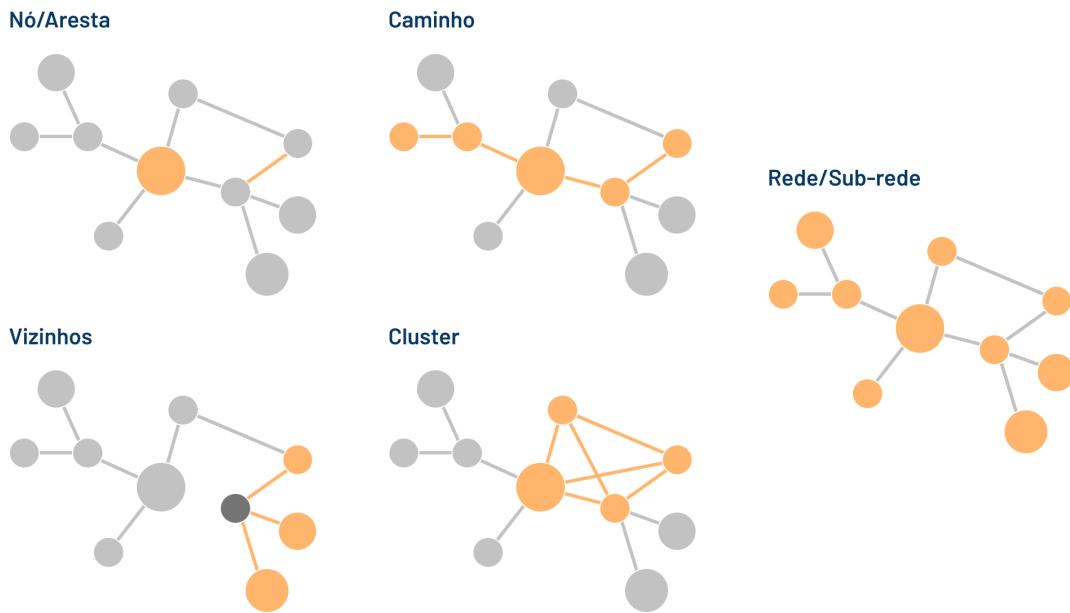
- **Analizar a topologia para determinados atributos**, (*analyzing Topology for given Attributes* (TgA)), que se destina à identificação, caracterização ou comparação de estruturas topológicas que possuem certos atributos.
- **Analizar atributos para uma determinada estrutura topológica**, (*analyzing Attributes for a given Topological structure* (AgT)), que visa identificar, caracterizar ou comparar os atributos de uma dada estrutura topológica.

Em Nobre *et al.* (2019), no contexto das tarefas de rede, os alvos são as estruturas topológicas, sendo considerados principais tipos: nós e arestas individuais, caminhos de nós e arestas sequenciais, vizinhanças (vizinhos) de um nó ou aresta, *cluster* de nós e arestas e redes ou sub-redes (Figura 6). Para Nobre *et al.* (2019) a estrutura topológica *cluster* refere-se a um conjunto de nós que estão bem conectados entre si, mas menos com outros nós no que diz respeito à topologia da rede. Por outro lado, a sub-rede é um subconjunto de uma rede que não está limitada a uma estrutura, ou seja, os nós não precisam estar bem conectados entre si.

Na Tabela 1, há exemplos das duas tarefas básicas (TgA e AgT) de análise de redes multivariadas aplicada ao conjunto de estruturas topológicas listadas anteriormente.

A Figura 7 apresenta um resumo e a relação entre todas as taxonomias de tarefas de

Figura 6 – Estruturas topológicas que são consideradas como alvos para tarefas por Nobre *et al.* (2019).



Fonte: Adaptada de Nobre *et al.* (2019).

Tabela 1 – Exemplos de tarefas de análise de redes multivariadas compostas por tarefas básicas e a estrutura topológica

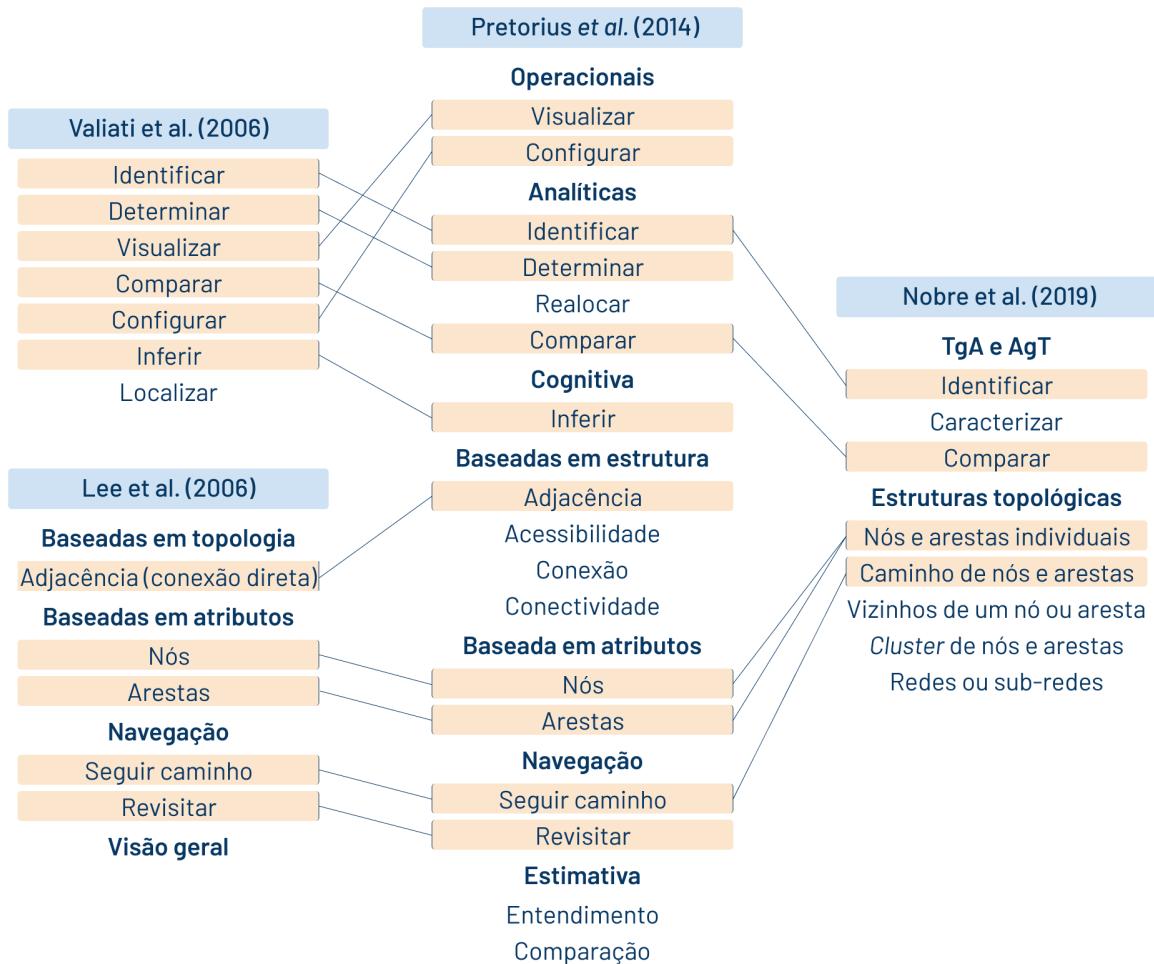
Estrutura Topológica	Tipo de tarefa	Exemplo
Nó/Link	TgA	Qual nó/link tem o valor de atributo fornecido?
	AgT	Qual é o valor do atributo de um determinado nó/link?
Vizinhos	TgA	Quais vizinhos têm um determinado valor de atributo?
	AgT	Quais são os valores de atributo dos vizinhos de um nó?
Caminhos	TgA	Qual caminho entre dois nós tem o menor custo?
	AgT	Quais são os valores de atributo ao longo de um caminho?
Clusters	TgA	Qual cluster tem determinadas características de atributo?
	AgT	Quais são as características de atributo dentro de um cluster?
Rede	TgA	Localize a sub-rede de nós com um atributo específico
	AgT	Quais são os valores de atributo nessa rede?

Fonte: Adaptada de Nobre *et al.* (2019).

análise de rede analisadas nesse capítulo. A partir da imagem pode-se observar que as tarefas que mais se repetem nas taxonomias são identificar e comparar, não aparecendo apenas na taxonomia de Lee *et al.* (2006). Os atributos que mais se repetem são nós e arestas, e o item de caminho em navegação, nas taxonomias de Lee *et al.* (2006) e Pretorius *et al.* (2014), e em estrutura topológica na taxonomia de Nobre *et al.* (2019).

O presente trabalho propõe que a avaliação da visualização de dados utilize tarefas de análise de rede com base em taxonomias, visto que as taxonomias de tarefas organizam as diferentes técnicas e métodos usados para estudar e compreender as redes.

Figura 7 – Resumo e relação entre as taxonomias de tarefas especializadas para visualização em redes apresentadas neste capítulo



Fonte: Elaborada pela autora

### 2.3 Avaliação de Interação Humano-Computador

O conceito de IHC originou-se com a invenção do primeiro dispositivo eletrônico de uso geral chamado *Electronic Numerical Integrator and Calculator* (ENIAC) em 1946 (SINHA *et al.*, 2010). Com a introdução de *Graphical User Interface* (GUI) e internet, percebeu-se a importância da IHC. Atualmente, a maioria dos softwares possui interface de usuário com a qual os usuários interagem, comunicando suas intenções, executando ações e recebendo *feedbacks*.

A IHC é uma disciplina preocupada com o projeto, avaliação e implementação de sistemas de computação interativos para uso humano e com o estudo dos principais fenômenos que os cercam (SINHA *et al.*, 2010). Tem como objetivos produzir sistemas usáveis, seguros e funcionais (ROCHA; BARANAUSKAS, 2003). Para isso, a IHC e a *User Experience* (UX) se beneficiam de conhecimentos e métodos de outras áreas fora da Computação para conhecer

melhor os fenômenos envolvidos no uso de sistemas computacionais interativos (BARBOSA *et al.*, 2021).

A avaliação de IHC é uma atividade fundamental em qualquer processo de desenvolvimento que busque produzir um sistema interativo com alta qualidade de uso (BARBOSA *et al.*, 2021). Não é suficiente apenas fornecer produtos com excelência técnica, os produtos também precisam ser fáceis de utilizar e se adaptar às práticas e atividades do usuário (BEVAN, 1999). A qualidade de uso é descrita na ISO/IEC 9126-1 (2001), como a visão do usuário sobre a qualidade do produto quando ele é usado em um ambiente específico e um contexto específico de uso. Os critérios de qualidade de uso enfatizam certas características da interação e da interface que as tornam adequadas aos efeitos esperados do uso do sistema (BARBOSA *et al.*, 2021). Os critérios de qualidade de uso, descritos por Barbosa *et al.* (2021) são: usabilidade, experiência do usuário, acessibilidade e comunicabilidade. A usabilidade enfoca a maneira como o uso de um sistema interativo é afetado por características do usuário. Essa qualidade relacionada aos sentimentos e emoções dos usuários é denominada experiência do usuário. O critério de acessibilidade está relacionado à remoção das barreiras que impedem mais usuários de serem capazes de acessar e interagir com um sistema (BARBOSA *et al.*, 2021). O critério de comunicabilidade chama a atenção para a responsabilidade do designer de comunicar ao usuário suas intenções de design e a lógica que rege o comportamento da interface (BARBOSA *et al.*, 2021).

Por meio da avaliação, o avaliador faz um julgamento de valor sobre a qualidade de uso do sistema, identificando problemas na interação e na interface que possam prejudicar a experiência do usuário. Avaliar o que foi construído é o foco do design da interação, o objetivo é assegurar que o produto seja adequado (PREECE *et al.*, 2005). A avaliação do produto final possibilita entregar um produto com uma garantia maior de qualidade, para que os problemas encontrados sejam corrigidos antes do produto chegar ao consumidor (BARBOSA *et al.*, 2021).

O papel da avaliação é garantir que o entendimento das necessidades dos usuários ocorra durante todos os estágios de desenvolvimento do produto (PREECE *et al.*, 2005). É importante que a solução de IHC seja avaliada do ponto de vista dos usuários, preferencialmente com a participação deles durante a avaliação (BARBOSA *et al.*, 2021), visto que a qualidade do sistema depende de como ele é representado e usado pelos usuários (KARRAY *et al.*, 2008).

A avaliação de IHC podem ser aplicada em diferentes momentos do processo de desenvolvimento. Quando a avaliação é realizada durante a elaboração da solução, ela é chamada de avaliação formativa ou construtiva. Já quando a avaliação é realizada depois de uma solução

estar pronta, ela é chamada de somativa ou conclusiva. A avaliação formativa é realizada ao longo de todo o processo de design para compreender e confirmar a compreensão sobre o que os usuários querem e precisam, e para confirmar se o grau em que a solução está sendo concebida atende às necessidades dos usuários (BARBOSA *et al.*, 2021), permitindo identificar problemas enquanto os custos de correção ainda são baixos. A avaliação formativa tem como objetivo orientar o design da solução para o caminho correto para que o produto tenha uma experiência de usuário satisfatória (JOYCE, 2019). A avaliação somativa é realizada ao final do processo de design, para julgar a qualidade de uso buscando evidências de que o produto possui os níveis de qualidade de uso desejados.

Várias técnicas estão disponíveis para fornecer suporte ao design e à avaliação (PRE-ECE *et al.*, 2005). Cada método atende melhor a certos objetivos de avaliação, orienta explícita ou implicitamente quando e onde os dados devem ser coletados, como eles devem ser analisados, e quais critérios de qualidade de uso (usabilidade, experiência do usuário, acessibilidade ou comunicabilidade) sua análise privilegia (BARBOSA *et al.*, 2021). Os métodos podem ser classificados em: investigação, observação e inspeção. Os métodos de investigação envolvem o uso de questionários, entrevistas, grupos focais e estudos de campo. Os métodos de inspeção permitem o avaliador examinar um sistema para antever os possíveis problemas que os usuários possam vir a ter. Alguns métodos de inspeção são: avaliação heurística, percurso cognitivo e a inspeção semiótica. Os métodos de observação fornecem dados sobre situações reais em que os usuários realizam suas atividades, permitindo identificar problemas reais, esses dados podem ser coletados por meio de: teste de usabilidade, método de avaliação de comunicabilidade e prototipação em papel.

Muitos estudos que comumente são realizados em laboratório podem ser facilmente conduzidos de forma remota, como entrevistas ou grupos focais (SCHMIDT *et al.*, 2021). Além disso, para observar os usuários durante as tarefas remotas, os recursos de compartilhamento de tela de diversas plataformas disponíveis podem ser usados (SCHMIDT *et al.*, 2021).

Os problemas na interação e na interface são avaliados por todos os métodos de avaliação citados anteriormente. Entretanto, para avaliar a forma como os usuários se apropriam dos sistemas computacionais interativos, utilizam-se entrevistas, estudos de campo, testes de usabilidade e avaliação da comunicabilidade. Em relação a produção de resultados quantitativos ou qualitativos, cada método costuma privilegiar um dos dois resultados (BARBOSA *et al.*, 2021).

Com a coleta e produção dos dados, há a atividade de interpretação, onde o avaliador analisa o material registrado para atribuir significado aos dados coletados. A interpretação do avaliador deve ser orientada pelo método de avaliação definido durante a atividade de preparação. Após a interpretação, os resultados são consolidados por meio da busca de recorrências nos dados. Os resultados de uma avaliação de IHC normalmente indicam tendências de problemas, e não uma certeza de que eles vão ocorrer durante o uso do sistema. Segundo Barbosa *et al.* (2021), os dados consolidados costumam incluir: os objetivos e escopo da avaliação, a forma como a avaliação foi realizada (qual método foi empregado), o número e o perfil dos participantes e avaliadores, um sumário dos dados coletados, um relato da interpretação e análise dos dados, uma lista de problemas encontrados e um planejamento para o reprojeto do sistema.

### **2.3.1 Usabilidade**

A *International Organization for Standardization* (ISO) define o conceito de usabilidade como: a medida em que um sistema, produto ou serviço pode ser usado por usuários específicos para atingir objetivos específicos com eficácia, eficiência e satisfação em um contexto específico de uso (ISO 9241, 2010). Segundo a norma, a eficácia está relacionada com a precisão e a integridade com que os usuários alcançam objetivos específicos. A eficiência está relacionada com os recursos gastos em relação à precisão e integridade com que os usuários alcançam seus objetivos. Já a satisfação está relacionada a desconfortos e atitudes positivas em relação ao uso do produto.

Nielsen *et al.* (2012) define a usabilidade como um atributo de qualidade que avalia o quanto fáceis de usar são as interfaces do usuário. A usabilidade de um sistema depende de quanto bem seus recursos acomodam as necessidades e os contextos dos usuários (Interaction Design Foundation - IxDF, 2016). Segundo Nielsen (1994), com tradução para o português de Barbosa *et al.* (2021), a usabilidade é definida por 5 fatores de qualidade:

- **Facilidade de aprendizado** (*learnability*): O sistema deve ser fácil de aprender para que o usuário possa rapidamente começar a trabalhar com o sistema.
- **Eficiência** (*efficiency*): O sistema deve ser eficiente de usar para que, uma vez que o usuário tenha aprendido o sistema, um alto nível de produtividade seja possível.
- **Facilidade de uso** (*memorability*): O sistema deve ser fácil de lembrar para que o usuário casual consiga retornar ao sistema após algum período sem utilizá-lo,

sem precisar aprender tudo de novo.

- **Segurança no uso (errors):** O sistema deve ter uma taxa de erro baixa para que os usuários cometam poucos erros durante o uso do sistema, e para que, se cometerem erros, possam se recuperar facilmente deles. Além disso, erros catastróficos<sup>6</sup> não devem ocorrer.
- **Satisfação (satisfaction):** O sistema deve ser agradável de usar para que os usuários fiquem subjetivamente satisfeitos ao utilizá-lo; eles gostem dele.

A usabilidade de um sistema é o alcance e o grau pelo qual o sistema pode ser usado de forma eficiente e adequada para atingir determinados objetivos para determinados usuários (KARRAY *et al.*, 2008). Normalmente, a usabilidade é medida em relação ao desempenho dos usuários em um determinado conjunto de tarefas (NIELSEN, 2001). As medidas mais básicas são: taxa de sucesso (se os usuários conseguem executar a tarefa), o tempo que uma tarefa requer para ser concluída, a taxa de erros e a satisfação subjetiva dos usuários. Também é possível coletar métricas mais específicas, como a porcentagem de tempo que os usuários seguem um caminho de navegação ou o número de vezes que eles precisam voltar atrás (NIELSEN, 2001).

Dados que abordam a satisfação e a percepção de usabilidade dos usuários são frequentemente coletados em estudos quantitativos junto com outros tipos de medidas de desempenho. Questionários populares são: *System Usability Scale* (SUS), pós-teste de usabilidade, que fornece informações úteis sobre as conclusões e a experiência geral de um usuário; *Single Ease Question* (SEQ), pós-tarefa, que oferece informações sobre a usabilidade de diferentes fluxos de tarefas; e NASA-TLX, pós-tarefa, que é apropriado para medir a carga de trabalho em tarefas complexas e críticas e *Usefulness, Satisfaction, and Ease of use* (USE) pós-teste, para medir a percepção dos usuários em relação à eficiência, eficácia e satisfação. Como todas essas são medidas quantitativas, elas exigem um tamanho de amostra razoavelmente grande para fornecer resultados válidos (LAUBHEIMER, 2018).

É importante conhecer as necessidades dos usuários e estabelecer quais critérios de usabilidade devem ser priorizados (BARBOSA *et al.*, 2021). A usabilidade não é uma propriedade única e unidimensional de uma interface de usuário (NIELSEN, 1994), se qualquer atributo de usabilidade falhar, a experiência geral do usuário será comprometida e muitos usuários falharão. É preciso ter uma visão integrada da usabilidade, da perspectiva do usuário, e desenvolver por meio de um design centrado no usuário (NIELSEN, 2008).

---

<sup>6</sup> Problemas que possam impedir que o usuário realize suas tarefas e alcance seus objetivos. Um erro catastrófico é um problema extremamente importante ser consertado antes de se lançar o produto.

Existem muitos métodos para estudar usabilidade, mas o mais básico e útil é o teste com usuários, que tem 3 componentes: usuários representativos, pedir aos usuários para executar tarefas representativas com o design e observar o que os usuários fazem, onde eles têm sucesso e onde eles têm dificuldades com a interface (NIELSEN *et al.*, 2012). Envolve a observação de usuários reais enquanto eles interagem com o sistema, o teste de usabilidade ajuda a identificar problemas na interface que podem não ser evidentes para os designers ou desenvolvedores.

### **2.3.2 Teste de usabilidade**

O teste de usabilidade é um método de avaliação por observação que visa avaliar a usabilidade de um sistema interativo a partir de experiências de uso dos seus usuários (RUBIN; CHISNELL, 2008). O teste de usabilidade dirá se o público-alvo de um produto pode usá-lo, ajuda a identificar problemas que as pessoas têm com uma interface específica e revela tarefas difíceis de concluir e linguagem confusa (GOODMAN *et al.*, 2012).

Os objetivos dos testes de usabilidade variam de acordo com o estudo, mas geralmente incluem: identificar problemas no design do produto ou serviço, descobrir oportunidades para melhorar e aprender sobre o comportamento e as preferências do usuário (MORAN, 2019). A intenção é garantir a criação de produtos que sejam úteis e valorizados pelo público-alvo, sejam fáceis de aprender, ajudem as pessoas a serem eficazes e eficientes naquilo que desejam fazer e sejam satisfatórios de usar (RUBIN; CHISNELL, 2008).

Os objetivos da avaliação determinam quais fatores de usabilidade devem ser medidos, por exemplo, caso o objetivo do estudo seja avaliar a facilidade de aprendizado, um teste de usabilidade poderá fornecer respostas a perguntas como: “quantos erros os usuários cometem nas primeiras sessões de uso?”, “quantos usuários conseguiram concluir com sucesso uma determinada tarefa?” e “quantas vezes os usuários consultaram a ajuda?” (BARBOSA *et al.*, 2021).

Os testes de usabilidade empregam técnicas para coletar dados empíricos enquanto observam usuários finais representativos usando o produto para executar tarefas realistas (RUBIN; CHISNELL, 2008). No teste de usabilidade, um grupo de usuários é convidado a realizar um conjunto de tarefas utilizando o sistema. Durante a interação com o sistema, as experiências de uso são observadas e são coletados dados sobre o desempenho dos participantes nas realizações das tarefas e suas opiniões sobre a experiência de uso.

De acordo com Goodman *et al.* (2012), existem quatro tipos principais de testes de

usabilidade:

- Exploratório, para testar conceitos preliminares.
- Avaliação, para testar recursos durante a implementação.
- Comparação, para avaliar um projeto em relação a outro.
- Validação, para certificar que os recursos atendem a certos padrões e *benchmarks*<sup>7</sup> no final do processo de desenvolvimento.

Normalmente, os testes de usabilidade são feitos como parte de um projeto de pesquisa maior e envolvem extensa preparação e análise (GOODMAN *et al.*, 2012). Para Rubin e Chisnell (2008) o processo para a condução de um teste de usabilidade conta com as etapas: desenvolver o plano de teste; configurar um ambiente de teste; encontrar e selecionar os participantes; preparar os materiais de teste; conduzir as sessões de teste; *debrief* com os participantes e observadores<sup>8</sup>; analisar os dados e observações; e relatório com resultados e recomendações. Para Goodman *et al.* (2012) as atividades de preparação do teste de usabilidade são: recrutar, escolher quais recursos testar, criar as tarefas e escrever um roteiro.

De acordo com Bastien (2010), a implementação de um teste de usuário geralmente passa por um certo número de etapas, tais como:

- Definir os objetivos;
- Qualificar e recrutar os participantes;
- Selecionar as tarefas que os participantes realizarão;
- Criar e descrever os cenários de tarefas;
- Escolher as medidas que serão coletadas, bem como a forma como os dados serão registrados;
- Preparar os materiais e o ambiente;
- Escolher o facilitador e criar o roteiro do teste;
- Criar e/ou selecionar os questionários de satisfação e os procedimentos de análise de dados
- Apresentar e comunicar os resultados.

A Tabela 2 apresenta as atividades do teste de usabilidade definidas por Barbosa *et al.* (2021).

---

<sup>7</sup> *Benchmark* é um processo de comparação sistemática de produtos, serviços, processos ou práticas de uma organização com os de outras. O objetivo é identificar áreas de melhoria, aprender com os melhores exemplos e implementar mudanças que possam aumentar a eficiência, qualidade e desempenho.

<sup>8</sup> O *debriefing* com os participantes refere-se a explorar e revisar as ações do participante durante o teste de usabilidade. O *debriefing* com os observadores consiste em fazer com que os avaliadores vejam diferentes pontos

Tabela 2 – Atividades do teste de usabilidade

Atividade	Tarefa
Preparação	Definir tarefas para os participantes executarem Definir o perfil dos participantes e recrutá-los Preparar material para observar e registrar o uso Executar um teste-piloto
Coleta de dados	Observar e registrar a performance e a opinião dos participantes durante sessões de uso controladas
Interpretação	Reunir, contabilizar e sumarizar os dados coletados dos participantes
Consolidação dos resultados	
Relato dos resultados	Relatar a performance e a opinião dos participantes

Fonte: Adaptada de Barbosa *et al.* (2021).

Os testes de usabilidade podem ser realizados em diferentes momentos do processo de desenvolvimento de um produto, desde protótipos iniciais até versões finais, e podem envolver variações, como qualitativo ou quantitativo, remoto ou pessoalmente.

Os testes de usabilidade podem ser apenas qualitativos ou quantitativos. Os testes de usabilidade qualitativos e quantitativos são métodos complementares que atendem a objetivos diferentes (BUDIU, 2017). Os testes qualitativos concentram-se na coleta de *insights*, descobertas e narrativas sobre como os usuários usam o produto ou serviço, sendo recomendado para identificar problemas na experiência do usuário (MORAN, 2019). É frequentemente usado de forma formativa para informar o processo de design e canalizá-lo na direção certa (BUDIU, 2017). Os testes quantitativos são baseados em um grande número de participantes (geralmente mais de 30) (BUDIU, 2017) e concentram-se na coleta de métricas que descrevem a experiência do usuário, as métricas mais comumente coletadas são o sucesso da tarefa e o tempo de execução (MORAN, 2019).

Em um teste de usabilidade remoto, os avaliadores e os participantes não estão no mesmo local físico, sendo uma solução para equipes com orçamento limitado ou para testar produtos cujos usuários estão geograficamente dispersos (SCHADE, 2013). De acordo com Bolt e Tulathimutte (2010), existem dois tipos principais de pesquisa remota: moderada e automatizada ou não-moderada.

O teste de usabilidade moderado remotamente é muito parecido com um estudo presencial, o facilitador ainda interage com o participante e pede que ele execute tarefa (MORAN, 2019). Na pesquisa moderada um facilitador conduzativamente os participantes por meio de uma série de tarefas, fazendo perguntas e monitorando suas reações ao produto, requer o mesmo tipo de engajamento que os testes presenciais, mas a interação humana ocorre *online* (GOODMAN *et al.*, 2012). A avaliação pode exigir aplicativos de videoconferência ou ferramentas de

---

de vista sobre o que está acontecendo nas sessões

compartilhamento de telas de modo a permitir que o avaliador veja o que está acontecendo na tela do usuário (BASTIEN, 2010).

Um estudo (HARTSON *et al.*, 1996), realizado em 1996<sup>9</sup>, demonstrou que o teste remoto moderado produz resultados comparáveis a um teste de usuário tradicional da mesma aplicação. Thompson *et al.* (2004) realizaram um estudo para identificar ferramentas apropriadas e definir metodologias para ambientes de testes remotos eficientes e eficazes, o estudo mostra que os testes remotos podem ser igualmente eficazes como testes tradicionais.

Os testes de usabilidade remotos não moderados não têm a mesma interação facilitadora-participante que os testes presenciais ou moderados (MORAN, 2019). Em vez disso, são configuradas instruções e tarefas para que sejam realizadas pelo participante. Para Whitenton (2019) as etapas para executar testes de usabilidade não moderados são: definir objetivos de estudo, selecionar software de teste, escrever as descrições das tarefas, realizar um teste piloto, recrutar participantes e analisar os resultados. Devido à falta de acompanhamento detalhado, é preferível utilizar testes de usabilidade remotos não moderados quando o foco principal do estudo são alguns elementos específicos, em vez de uma revisão geral (SCHADE, 2013).

Nelson e Stavrou (2011) fizeram testes de usabilidade assíncronos remotos (não moderados) fazendo com que os usuários resolvessem tarefas e avaliassem a usabilidade, em seguida, os avaliadores analisaram *logs* de cliques e pressionamentos de teclas dos usuários. Nessa configuração, os problemas de usabilidade foram identificados pelos avaliadores, e não pelos usuários, embora com base em dados restritos.

Resumidamente, o teste de usabilidade é um método de avaliação de IHC que visa avaliar a usabilidade a partir da experiência de uso dos usuários, podendo ser realizados em diferentes momentos do processo de desenvolvimento e possuem variações, como: qualitativo ou quantitativo, remoto ou presencial, moderado ou não-moderado. Ademais o método permite a combinação de métodos, como: entrevistas e questionários, para fornecer uma visão mais abrangente.

## 2.4 Considerações Finais

Neste capítulo, apresentaram-se os conceitos necessários para a compreensão deste trabalho: uma introdução à visualização de dados, uma visão geral sobre as abordagens e as

---

<sup>9</sup> O artigo discute os benefícios dos testes de usabilidade remotos, que permitem testes mais frequentes e diversos do que os métodos tradicionais baseados em laboratório.

taxonomias de tarefas de análise de redes multivariadas, e os métodos de avaliação de IHC.

Observou-se que as redes multivariadas são muito utilizadas e fornecem desafios particulares para a comunidade de pesquisa de visualização de dados (KERREN *et al.*, 2014). Para que aplicações envolvendo redes multivariadas atendam às necessidades dos usuários e forneçam informações relevantes de forma clara e acessível, é fundamental uma metodologia de avaliação que se baseie não só nas taxonomias de tarefas de redes, já consolidadas na comunidade de visualização, mas também em metodologias de avaliação consolidadas na área de IHC. A abordagem proposta neste trabalho pretende atender a essa necessidade, utilizando o método de observação de uso teste de usabilidade com a inclusão de tarefas de análise de rede.

No próximo capítulo são apresentados os trabalhos relacionados à avaliação de visualização de dados e de redes multivariadas.

### **3 TRABALHOS RELACIONADOS**

Os trabalhos apresentados neste capítulo têm como tema principal a avaliação de visualizações de dados e de redes multivariadas, com o intuito de auxiliar na tomada de decisões metodológicas e de desenvolvimento da metodologia para avaliação com usuários de aplicações baseadas na visualização de redes multivariadas. Esses trabalhos relacionados foram utilizados como base para a realização do presente trabalho.

A busca por trabalhos relacionados se deu pelo método *Quick and Dirty Review* (YI, 2014), que consiste em identificar, em um curto período de tempo, o que ler com mais relevância antes de investir tempo. Os artigos foram encontrados pela ferramenta Google Scholar<sup>10</sup>, utilizando-se as palavras chaves: visualização de dados, rede multivariada e avaliação, em inglês. Após encontrados os artigos e realizada uma breve leitura conforme recomendada pelo método, construiu-se uma planilha com as informações: referência, escala de interesse, ideia central, novas ideias que o artigo desperta e o grau de excelência da publicação (Qualis). Com a planilha concluída com 100 artigos, foi feita uma classificação dos artigos de acordo com o nível de interesse, para que fosse realizada uma leitura completa e detalhada dos artigos de maior interesse para este trabalho.

#### **3.1 Métodos de avaliação de visualização de dados**

Nesta seção, serão abordados trabalhos sobre os métodos de avaliação de visualização de dados mais utilizados ou que propõem metodologias para a avaliação, e as principais características relacionadas.

Em um levantamento da literatura com cerca de 50 estudos com usuários de sistemas de visualização de dados, Plaisant (2004) encontrou quatro temáticas de avaliação: experimentos controlados comparando elementos de design, avaliação de usabilidade, experimentos controlados comparando duas ou mais ferramentas e estudos de casos. Sedlmair *et al.* (2012) afirmam que a forma mais comum de validação de visualização de dados são estudos de casos com usuários reais, problemas reais e dados reais.

Carpendale (2008) afirma que as visualizações de dados são frequentemente avaliadas usando pequenos conjuntos de dados, com participantes universitários e usando tarefas simples. Para incentivar o interesse de potenciais usuários, as visualizações de dados precisam ser testadas

---

<sup>10</sup> <https://scholar.google.com/>

com usuários reais, tarefas reais e também com conjuntos de dados grandes e complexos (CARPENDALE, 2008).

É importante destacar que o termo “usuários reais”, aqui apresentado, refere-se a usuário potencial (quem tem o perfil ou as necessidades que indicam que poderia se beneficiar da solução) ou usuário final (quem utiliza diretamente a solução), visto que normalmente as visualizações de dados são avaliadas com estudantes universitários, usuários casuais (pessoas que utilizam a solução de forma esporádica e sem profundo conhecimento técnico de análise de dados) ou com especialistas em visualização.

Em uma outra análise de estudos com usuários em uma revisão da literatura na área de visualização de informação, Ellis e Dix (2006) relatam que, de 65 artigos descrevendo novas aplicações ou técnicas de visualização, apenas 12 mencionaram terem feito algum tipo de avaliação com usuário, desses 12, vários eram problemáticos e apenas dois estudos foram considerados bem sucedidos. A maioria dos artigos tenta justificar a relevância da técnica proposta por meio de exemplos, argumentando que esta é superior a outras, mas que o uso de comparações diretas não é comum.

Os estudos de caso podem descrever todo o processo de descoberta, colaborações entre usuários, as frustrações da limpeza de dados e a emoção de ver a primeira visão geral dos dados (PLAISANT, 2004). Em geral, os estudos de caso são a estratégia preferida quando as questões de pesquisa do tipo “como” ou “por quê” são utilizadas, quando o investigador tem pouco controle sobre os acontecimentos e quando o foco está em um fenômeno contemporâneo dentro de algum contexto da vida real (YIN, 2014).

Após uma revisão histórica dos métodos de avaliação, Shneiderman e Plaisant (2006) descrevem um método de pesquisa chamado *Multi-dimensional In-depth Long-term Case studies* (MILCs), que foi desenvolvido para avaliar ferramentas de apoio à criatividade, e pode ser usado para avaliar ferramentas de visualização de informação. O MILCs baseia-se na noção de estudos de campo ou de caso utilizando métodos etnográficos de observação dos participantes, além de entrevistas, pesquisas e registro automatizado da atividade do usuário (SHNEIDERMAN; PLAISANT, 2006).

Para tornar a avaliação uma parte significativa das práticas de visualização de dados, Isenberg *et al.* (2008) defendem o método “*grounded evaluation*”, concentrando-se em abordagens qualitativas, incluindo observação como parte do processo de design, entrevistas e estudos de campo, fornecendo dados reais e considerando a interação um fator que influencia as

visualizações, seu desenvolvimento e seu uso.

Executar avaliações com conjuntos de dados completos, tarefas específicas de domínio e especialistas de domínio como participantes ajudará a desenvolver evidências muito mais concretas e realistas da eficácia de uma determinada visualização de dados (CARPENDALE, 2008).

Huang *et al.* (2009) introduzem duas medidas para avaliação de visualizações: esforço mental e eficiência da visualização, propondo um modelo cognitivo para revelar as relações interativas entre desempenho da tarefa, esforço mental dedicado durante o desempenho e demanda de memória induzida pela tarefa.

Nobre *et al.* (2021) apresentam uma nova metodologia de análise para avaliar visualizações interativas complexas, fornecendo um conjunto de diretrizes aplicáveis tanto a métodos quantitativos quanto qualitativos. Ao avaliar visualizações interativas complexas com estudos empíricos, os métodos de análise tradicionais não levam em conta variações nas estratégias de análise de participantes. Para enfrentar esse desafio, Nobre *et al.* (2021) desenvolveram o sistema *open-source* ReVISit. O ReVISit apresenta um fluxo de trabalho e um conjunto de métodos para capturar e analisar dados detalhados para esclarecer como as estratégias afetam os resultados do estudo. Para isso, é recomendável capturar interações e registrar informações sobre o elemento da interface do usuário com o qual ele interagiu.

Vários métodos de avaliação de visualização propõem o registro das atividades do usuário com a visualização ou observação da interação (SHNEIDERMAN; PLAISANT, 2006; ISENBERG *et al.*, 2008; NOBRE *et al.*, 2021).

Dentre as características apresentadas anteriormente, os aspectos que serão norteadores deste trabalho são: avaliação com usuários reais, dados reais e a utilização de tarefas específicas de domínio. Além disso, considera-se a interação como um fator que influencia o uso das visualizações de dados.

A abordagem de avaliação proposta neste trabalho, assim como defendido por Carpendale (2008), parte do princípio de que a avaliação deve possuir tarefas específicas de domínio e especialistas de domínio como participantes, mas também está baseada nas taxonomias de tarefas de rede para elaboração das tarefas de domínio, permitindo uma avaliação mais objetiva e direcionada da visualização. Além disso, a abordagem proposta também combina características metodológicas, já consolidadas em IHC, levando em consideração outros aspectos de usabilidade além da eficiência e eficácia.

### **3.2 Avaliação de visualização de redes multivariadas**

Vários estudos concentram-se na comparação de representações visuais de redes, comparando as representações nó-link e matriz de adjacência (GHONIEM *et al.*, 2005; KELLER *et al.*, 2006; JIANU *et al.*, 2014; OKOE *et al.*, 2018; REN *et al.*, 2019; NOBRE *et al.*, 2020; ABDELAAL *et al.*, 2022). A pesquisa existente em estudos de avaliação de rede até agora tem focado principalmente na topologia da rede (NOBRE *et al.*, 2020). Muitos estudos que compararam os tipos de visualização de redes geralmente se concentram na eficiência dos algoritmos de layout ou nas diferenças estéticas (REN *et al.*, 2019). Nesta seção, serão apresentados trabalhos que avaliam a visualização de redes multivariadas com usuários.

Ghoniem *et al.* (2005) avaliaram a legibilidade das abordagens representações nó-link e matriz de adjacência usando sete tarefas genéricas baseadas em topologia em grafos com diferentes tamanhos (20, 50 e 100 nós) e diferentes densidades de arestas (20%, 40% e 60%) para cada tamanho, com um total de nove gráficos diferentes. No estudo, a legibilidade é definida como a facilidade em que o usuário encontra a informação que ele está procurando, sendo consideradas com mais legibilidade visualizações em que o usuário executa tarefas mais rápido e comete menos erros. O estudo foi realizado com 36 participantes, sendo composto por alunos de pós-graduação e pesquisadores da área da Ciência da Computação habituados com representações nó-link. A interatividade limitava-se à seleção de nós e arestas. Descobriu-se que diagramas nó-link superam matrizes de adjacência em gráficos pequenos e esparsos, mas para gráficos maiores ou mais densos, matrizes de adjacência produziram resultados mais precisos. A exceção são as tarefas baseadas em caminho, em que os diagramas de nó-link superam a matriz de adjacência independentemente do tamanho ou densidade da rede.

De forma similar, Keller *et al.* (2006) relatam os resultados de dois estudos de usuários sobre a legibilidade. O primeiro foi conduzido para identificar os principais atributos que influenciam a legibilidade de representações matriciais de gráficos. O segundo, examina qual representação é mais adequada para exibir certos aspectos da estrutura relacional. O primeiro experimento foi realizado com 21 participantes, compostos por estudantes de doutorado ou profissionais de engenharia, o experimento foi conduzido de forma remota. O segundo, foi realizado com 16 participantes, também com estudantes de doutorado ou profissionais de engenharia, usando seis tarefas em uma rede direcionada de domínio específico. Os resultados se correlacionaram com os descritos em Ghoniem *et al.* (2005).

Muitos estudos foram conduzidos com um grupo maior de participantes, por meio

da ferramenta de *Crowdsourcing* Amazon Mechanical Turk (MTurk)<sup>11</sup>.

Diferente da comparação entre as duas abordagens de visualização de rede, Jianu *et al.* (2014) avaliaram quatro técnicas para exibir informações de grupo ou *cluster* sobrepostos no diagrama nó-link: coloração de nós, GMap<sup>12</sup>, BubbleSets<sup>13</sup> e LineSets<sup>14</sup>, em um estudo on-line utilizando o MTurk. Foram avaliados 10 tipos diferentes de tarefas que se enquadram nas três categorias: tarefas de grupo, tarefas de rede e tarefas mistas de grupo-rede, inspiradas pela taxonomia de tarefas de grafos descrita em Lee *et al.* (2006). O estudo foi realizado com 788 usuários que inicialmente tiveram uma breve introdução ao objetivo do estudo, aos dados e às visualizações usadas, após isso foi permitido experimentar a visualização e seus recursos por 1 minuto. Foram coletados dados quantitativos e feito um estudo utilizando *eye-tracking*.

Okoe *et al.* (2018) apresentam os resultados de um estudo comparando a eficácia entre diagramas nó-link e matrizes de adjacência, utilizando dois *datasets*, com densidades diferentes, representativos de redes do mundo real em tamanho e estrutura. O estudo foi realizado com 864 usuários, contendo 14 tarefas, utilizando MTurk. As tarefas testaram a habilidade dos participantes para identificar e comparar grupos de nós ou *clusters* e memorização. As tarefas eram apresentadas com um limite de tempo, ocultando a visualização quando o tempo estivesse expirado. No final do estudo, era pedido aos participantes que fizessem comentários. No estudo foram coletados dados tanto qualitativos quanto quantitativos. Como resultado, notou-se que os diagramas nó-link possuem maior precisão e tempo de resposta mais rápido, mas a diferença do resultado em relação as matrizes de adjacência diminuiu à medida que os participantes se familiarizaram com as visualizações.

Ren *et al.* (2019) investigam o entendimento humano de diferentes visualizações de rede: diagrama nó-link e duas variações de ordenação de matriz para redes de 20 e 50 nós. O entendimento foi quantificado usando tempo de tarefa e métricas de acurácia em perguntas que foram derivadas das tarefas comuns na análise de grafos, desenvolvida por Lee *et al.* (2006), com perguntas que se encaixam em uma das cinco categorias de tarefas: Adjacência, Acessibilidade, Conexão Comum, Conectividade (caminho mais curto) e Atributo. Além disso, foi avaliado o aprendizado com várias perguntas gerais apresentadas em um questionário pós-estudo. O

<sup>11</sup> <https://www.mturk.com/>

<sup>12</sup> GMap, introduzido por Gansner *et al.* (2010), usa uma metáfora de mapa de preenchimento de espaço para delimitar os membros do grupo em “países”, “mares” e “lagos”.

<sup>13</sup> BubbleSets, introduzido por Collins *et al.* (2009), desenha contornos contíguos ao redor nós do mesmo grupo, mesmo que os nós não sejam espacialmente co-localizado.

<sup>14</sup> LineSets, introduzido por Alper *et al.* (2011), liga todos os membros de um conjunto com uma curva contínua de cor distinta.

estudo foi realizado com 600 participantes na plataforma MTurk, contendo 4 fases principais: questionário pré-teste, treinamento, tarefas e questionário pós-teste. Foram coletados não apenas os dados de desempenho de tarefas, mas também dados sobre interação com a interface (como o movimento do mouse e anotações dos movimentos). Os resultados indicaram que a visualização nó-link produz uma melhor compreensão dos dados, com maior precisão de resposta e tempo de conclusão das tarefas mais rápido. Com a coleta dos movimentos do mouse dos participantes, notou-se que a visualização de diagrama nó-link levou a velocidades de movimento significativamente mais altas durante as respostas às perguntas.

Nobre *et al.* (2020), em um experimento *crowdsourcing*, fizeram a comparação entre duas redes multivariadas: diagramas nó-link com codificação no nó e matrizes de adjacência com tabelas justapostas. Foram elaboradas oito hipóteses que foram avaliadas em um conjunto de 16 tarefas, formuladas a partir da análise de tarefas de Nobre *et al.* (2019), com 322 participantes. O procedimento consistia em cinco fases: treinamento passivo, treinamento ativo, testes, estudo, e dados demográficos e *feedback*. Ao longo do estudo, foram coletados dados qualitativos e quantitativos. Foi capturado o horário de início e término para cada fase, bem como o tempo navegado fora da janela do estudo, o que permitiu avaliar o tempo médio gasto com treinos e *trials*. Durante a fase do estudo, foram coletados o tempo gasto em cada tarefa, as respostas, a confiança na resposta apresentada e a dificuldade percebida da tarefa; os dois últimos em uma escala Likert de 7 pontos. Por meio das perguntas de resposta livre, coletou-se *feedback* qualitativo para cada tarefa. Na fase de dados demográficos e *feedback*, um formulário incluía perguntas de resposta livre em que os usuários forneceram *feedback* sobre o material de treinamento e sobre o estudo. Além disso, foram coletados dados valiosos das interações dos usuários, incluindo procurar por nó, arrastar um nó, (des)selecionar um nó, limpar nós selecionados, operações de classificação e passar o mouse em um nó. A Tabela 3 apresenta as 16 tarefas realizadas pelos participantes, abrangendo as principais estruturas topológicas: nós únicos, vizinhos, *clusters* e caminhos. Como resultado, descobriu-se que os diagramas nó-link são mais adequados para tarefas que exigem uma integração entre a topologia de rede (tarefas de caminhos, tarefas que dependem de comparações dentro do nó e tarefas de vizinhos) e alguns atributos. As matrizes de adjacência funcionam bem para tarefas relacionadas a *clusters* e quando mais atributos precisam ser considerados. Além disso, o estudo destacou a importância de treinamento e remuneração<sup>15</sup>

<sup>15</sup> No Brasil, o Conselho Nacional de Saúde (CNS) regulamenta as pesquisas científicas envolvendo pessoas, na Resolução No 466/2012 não permite remuneração para os participantes, como ocorre em outros países, podendo, apenas, receber ressarcimento de despesas efetuadas com transporte, hospedagem e alimentação.

para garantir participantes motivados.

Tabela 3 – Lista de tarefas e instruções dadas aos participantes

Nome da tarefa	Instrução
Pesquisa de nó por atributo	Encontre o norte-americano com mais <i>tweets</i> .
Pesquisa de nó por atributo com distratores	Encontre a pessoa ou instituição europeia com menos <i>likes</i>
Pesquisa de nó por topologia com múltiplos atributos	Qual pessoa tem muitas interações nessa rede, vários seguidores, mas poucos <i>tweets</i> e curtidas em geral?
Pesquisa de vizinho por atributo	Encontre todos os vizinhos europeus de Lane.
Pesquisa de vizinho por atributo com distratores	Encontre todos os vizinhos norte-americanos da giCentre
Pesquisa de vizinho por atributo de link	Quem teve mais interações do tipo menção com Jeffrey?
Visão geral de vizinho por atributo de link	Alex tem mais interações do tipo menção com contas norte-americanas ou europeias? Com quem ele tem o mais interações do tipo menção?
Atributo de vizinhos comuns	Entre todas as pessoas que interagiram com Jeffrey e Robert, quem tem mais seguidores?
Atributo de link	Qual é a forma mais comum de interação entre Evis19 e Jon? Quantas vezes isso aconteceu?
Comparação entre atributos de nó	Selecione todos os vizinhos de Noeska que são pessoas e têm mais amigos do que seguidores.
Comparação entre atributos de nó (rede pequena)	Selecione as pessoas que interagiram com Thomas e têm mais amigos do que seguidores.
Estimativa de <i>cluster</i> e atributo	Selecione todas as pessoas que estão em um cluster com Alex e estime sua média de seguidores.
Atributo no caminho mais curto	Qual é a instituição em um caminho mais curto entre Lane e Rob? Qual é o seu continente de origem?
Atributo no caminho mais curto (rede pequena)	Qual é a instituição em um caminho mais curto entre Jason e Jon? Qual é o seu continente de origem?
Atributo em múltiplos caminhos	Dos norte-americanos que estão a duas interações de distância de Sereno, quem está há mais tempo no Twitter?
Exploração livre	Explore a rede livremente e relate suas descobertas. Há algo surpreendente ou interessante?

Fonte: Nobre *et al.* (2020).

Em Abdelaal *et al.* (2022), foi realizada a comparação entre grafo bipartido<sup>16</sup> e as duas abordagens mais comuns para redes estáticas: diagramas nó-link e matrizes de adjacências. Foi realizado um estudo de usuários com 150 participantes, utilizando a plataforma MTurk, para avaliar o desempenho em 5 tarefas de interpretação de rede, sendo 3 de visão geral da rede: identificação de classe de rede, detecção de *cluster* e estimativa de densidade de rede, e duas tarefas detalhadas: grau de entrada vs. grau de saída do nó e mapeamento de representação. A maioria das tarefas focava na visão geral da rede, uma vez que, segundo os autores, muitas vezes ela não foi considerada por outros estudos. O estudo não fornecia uma interação com a visualização de dados, eram apenas mostradas ao participante imagens das abordagens de rede e feitas perguntas sobre uma técnica de visualização. O estudo destacou que o desempenho

<sup>16</sup> Grafos bipartidos são um tipo especial de grafo onde os vértices são divididos em dois conjuntos disjuntos.

de algumas tarefas poderiam ter sido melhor se tivessem sido realizadas em um laboratório controlado.

Dentre as características apresentadas anteriormente e na Seção 3.1, os aspectos que serão norteadores deste trabalho são: avaliação com usuários reais, utilização de visualização real, utilização de taxonomia de tarefas e avaliação da interação. Além disso, a abordagem proposta pode ser utilizada para comparar abordagens de redes multivariadas ou técnicas de visualização de dados. A Tabela 4 relaciona as características das avaliações apresentadas anteriormente.

Tabela 4 – Relação das características das avaliações

	Usuários reais	Visualizações reais	Taxonomia de tarefas	Avaliação da interação	Comparação entre abordagens
Ghoniem <i>et al.</i> (2005)					x
Keller <i>et al.</i> (2006)					x
Jianu <i>et al.</i> (2014)	x		x	x	
Okoe <i>et al.</i> (2018)	x				x
Ren <i>et al.</i> (2019)	x		x	x	x
Nobre <i>et al.</i> (2020)	x		x	x	x
Abdelaal <i>et al.</i> (2022)					x
Abordagem proposta	x	x	x	x	x

Fonte: Elaborada pelo autora.

Poucos trabalhos avaliam a interatividade existente e utilizaram alguma taxonomia de tarefas. São menos comuns os experimentos que avaliam o impacto do tipo de visualização no desempenho humano (REN *et al.*, 2019). Os trabalhos apresentados focam principalmente na eficiência (avaliando se o usuário consegue concluir com sucesso uma determinada tarefa), e eficácia (avaliando quanto tempo é necessário para o usuário concluir uma tarefa) (GHONIEM *et al.*, 2005; KELLER *et al.*, 2006; JIANU *et al.*, 2014; OKOE *et al.*, 2018; REN *et al.*, 2019; NOBRE *et al.*, 2020; ABDELAAL *et al.*, 2022).

Nenhum dos trabalhos realizou a avaliação com usuários reais. Como visto no capítulo anterior, para incentivar o interesse de potenciais usuários, as visualizações de dados precisam ser testadas com usuários reais e com conjuntos de dados grandes e complexos (CARPENDALE, 2008).

Dependendo do processo de geração, redes aleatórias muitas vezes não exibem características comuns a redes do mundo real (REN *et al.*, 2019). Ghoniem *et al.* (2005) e Keller *et al.* (2006) usam um único gráfico ou um conjunto de gráficos gerados aleatoriamente. Em Abdelaal *et al.* (2022) o estudo não fornecia uma visualização de dados, eram apenas mostradas ao participante imagens das abordagens de rede. Abdelaal *et al.* (2022), Nobre *et al.* (2020), Ren

*et al.* (2019) e Okoe *et al.* (2018), forneceram aos usuários uma fase de treinamento, antes de começar os estudos com usuários.

Ao usar dados reais, as avaliações permitem que as soluções, produtos ou modelos sejam testados em condições que refletem a realidade. Para avaliar se a visualização está fazendo sentido para o usuário, se os dados não forem reais eles precisam fazer parte do contexto do usuário, visto que um resultado é considerado realista na medida em que o contexto em que foi estudado é semelhante ao contexto em que será usado (CARPENDALE, 2008).

A abordagem de avaliação proposta utiliza o método de observação de uso teste de usabilidade e não é focada apenas em experimentos controlados apenas comparando abordagens de redes multivariadas (diagrama nó-link e matriz de adjacência) ou técnicas de visualização, como os trabalhados apresentados neste capítulo. No entanto, caso o objetivo da avaliação demande um experimento controlado para comparar duas abordagens diferentes, a abordagem de avaliação aqui proposta contempla a realização de desse tipo de experimento.

### 3.3 Considerações Finais

Neste capítulo, foram descritas abordagens de avaliação de visualizações de dados e de redes multivariadas.

Na Seção 3.1, observou-se que vários métodos de avaliação de visualização propõem o registro das atividades do usuário com a visualização ou observação da interação (SHNEIDERMAN; PLAISANT, 2006; ISENBERG *et al.*, 2008; NOBRE *et al.*, 2021), mas poucos utilizam tarefas específicas de domínio e especialistas de domínio como participantes (CARPENDALE, 2008).

Na Seção 3.2, constatou-se que poucos trabalhos avaliam a interatividade existente e utilizam alguma taxonomia de tarefas de análise, e nenhum realizou a avaliação com usuário reais. Essa etapa, foi essencial para o entendimento das características comumente avaliadas e definição dos aspectos considerados norteadores deste trabalho: avaliação com usuários reais, utilização de visualização reais, utilização de taxonomia de tarefas de análise de rede e avaliação da interação.

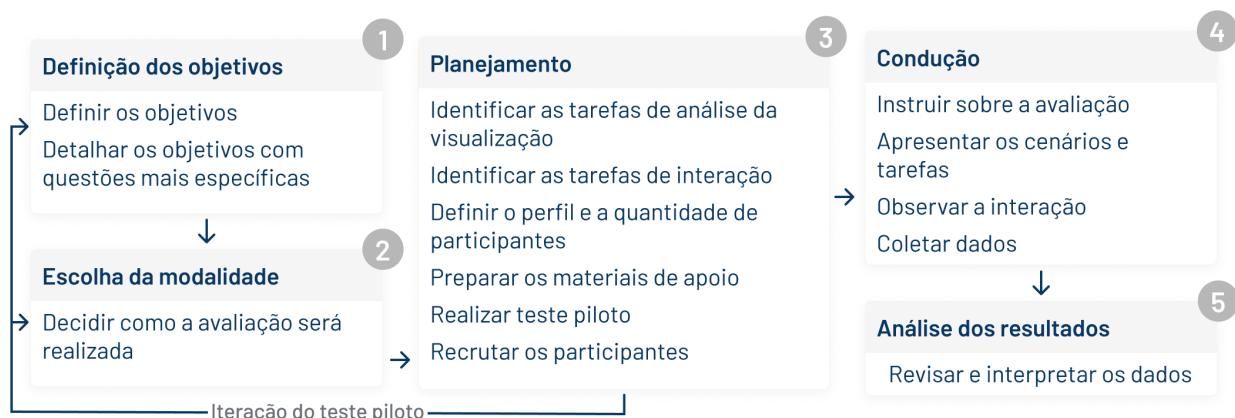
No capítulo seguinte, será apresentada a metodologia de avaliação proposta.

## 4 ADAPTAÇÃO DE TESTES DE USABILIDADE

Neste capítulo, será descrita a abordagem proposta para avaliar a visualização de redes multivariadas, envolvendo usuários, dados e tarefas reais, fundamentadas em uma taxonomia de tarefas, e levando em consideração fatores de usabilidade além da eficiência e eficácia. Para isso, é utilizado o método de avaliação de IHC: o método de observação de uso teste de usabilidade, com a inclusão de tarefas de análise de rede, e métodos de investigação entrevista ou questionário, no início e final da avaliação.

A sugestão é que a avaliação da visualização seja estruturada em cinco etapas: definição dos objetivos da avaliação, escolha da modalidade do teste de acordo com os objetivos, planejamento, condução e análise dos resultados. A Figura 8 apresenta um resumo das etapas e as seções seguintes descrevem em detalhes cada etapa.

Figura 8 – Fluxo de execução da abordagem proposta para avaliação de ferramenta de visualização de rede



Fonte: Elaborado pela autora.

### 4.1 Definição dos objetivos da avaliação

Os objetivos devem ser definidos com base em requisições, reclamações e ou comportamentos dos *stakeholders*<sup>17</sup> do sistema (BARBOSA *et al.*, 2021). Todo o planejamento, condução e apresentação dos resultados são orientados por esses objetivos (SHARPE *et al.*, 2007).

Para cada objetivo definido, o avaliador deve elaborar perguntas específicas que a avaliação deverá responder.

<sup>17</sup> Stakeholders são partes interessadas em uma organização, projeto ou iniciativa que têm influência direta ou indireta sobre ela

## 4.2 Escolha da modalidade de avaliação

Nessa etapa, há a escolha de como a avaliação será realizada, o avaliador deve decidir se a avaliação será presencial ou remota, moderada ou não-moderada e definir as fases que serão realizadas.

A avaliação pode abranger as fases: entrevista ou questionário pré-teste, exploração livre, observação de uso das tarefas e entrevista ou questionário pós-teste. O avaliador deve escolher as fases de acordo com os objetivos da avaliação.

No início da avaliação, a entrevista ou o questionário pré-teste deve ser utilizado para coletar dados sobre os participantes como: informações demográficas, contexto de uso, experiência prévia, motivação e expectativas, entre outros.

Durante a avaliação, os participantes serão convidados a realizar um conjunto de tarefas de análise de rede e de interação com o sistema. É recomendável a etapa de exploração livre, onde o participante explora o sistema livremente relatando suas descobertas e percepções por meio da técnica *think aloud*. Na técnica *think aloud* é solicitando aos participantes que usem o sistema enquanto pensam em voz alta — isto é, verbalizando seus pensamentos à medida que avançam pela interface (NIELSEN, 2012).

Na sessão de observação de uso, as tarefas são solicitadas ao participante em forma de cenário. Os cenários orientam os participantes sobre as tarefas a serem realizadas (BARBOSA *et al.*, 2021), situando a solicitação da tarefa em um cenário curto que prepare o participante para a ação e forneça uma explicação e contexto sobre por que ele está realizando a tarefa (MCCLOSKEY, 2014).

Nesse método, os participantes são apresentados a cenários realistas e contextuais que simulam situações práticas de uso de um sistema, produto ou interface. O objetivo é que o usuário realize uma tarefa específica dentro do contexto do cenário fornecido, o que gera um teste mais natural e fiel à realidade.

No fim, a entrevista ou o questionário pós-teste é utilizado para coletar a opinião do participante sobre a experiência de uso que acabou de vivenciar e esclarecer eventuais dúvidas sobre seu comportamento e suas intenções na execução das tarefas (BARBOSA *et al.*, 2021), além da coleta de *feedbacks* e sugestões. Questionários específicos, como: SUS, SEQ e NASA-TLX (seção 2.3.1) para coleta de dados da satisfação e a percepção de usabilidade dos usuários também podem ser empregados.

## 4.3 Planejamento

A etapa de planejamento é composta das seguintes atividades: identificar as tarefas de análise da visualização, identificar as tarefas de interação, definir o perfil e a quantidade de participantes, preparar os materiais de apoio, executar o teste piloto e recrutar os participantes. Cada uma dessas atividades é descrita a seguir.

### 4.3.1 Identificar as tarefas de análise da rede

Especialistas do domínio da aplicação devem enumerar as tarefas no contexto do domínio da aplicação que os usuários devem ser capazes de realizar com a visualização, de acordo com o objetivo da avaliação. A partir dessas tarefas, são identificadas as tarefas analíticas de rede correspondentes nas taxonomias de tarefas existentes.

As taxonomias de tarefas de redes existentes foram analisadas na Seção 2.2.3. Sugere-se escolher a taxonomia de Nobre *et al.* (2019), já que é a mais recente e construída a partir das taxonomias anteriores, e portanto a mais completa.

Dessa forma para cada tarefa de domínio, é encontrada uma ou mais tarefas de redes associadas, conforme listadas na Tabela 1. Essa correspondência pode ser feita com a ajuda de um especialista em visualização e é fundamental para avaliar se a técnica de visualização de redes utilizada atende às tarefas de domínio.

Por exemplo, no contexto de uma aplicação com a visualização de uma rede social, o especialista de domínio ou *stakeholder* da ferramenta afirma que o usuário deve ser capaz de encontrar as pessoas que interagiram com Maria no mês de julho. A partir dessa informação, consultando a taxonomia de Nobre *et al.* (2019), descobre-se que ela corresponderá a duas tarefas de rede: a primeira de identificação de um nó (alvo) com um dado atributo nome = ‘Maria’ (primeiro exemplo da primeira linha da Tabela 1) e em seguida uma tarefa de identificação que tem como alvo a estrutura topológica de vizinhos com atributo data\_mensagem.mês = ‘Julho’ (primeiro exemplo da segunda linha da Tabela 1).

É importante ressaltar que a prioridade na definição das tarefas de rede deve ser atender as tarefas de domínio, mesmo que estas correspondam a apenas um subconjunto das tarefas da taxonomia adotada. O domínio deve ser traduzido em tarefas de usuário e as tarefas de usuário devem ter relação direta com a visualização de dados projetada. Se a técnica de visualização não atender a tarefa de domínio (e de rede) proposta, essa tarefa deve ser excluída

da avaliação. Da mesma forma, tarefas de rede sem uma tarefa de domínio associada não devem ser inseridas na avaliação pois elas não contribuem para os objetivos da aplicação.

A compreensão da tarefa baseada em taxonomia permite a avaliação da solução de visualização para resolver o problema de domínio e a comparação posterior com outras aplicações com visualização de redes que estejam em um domínio distinto.

#### ***4.3.2 Identificar as tarefas de interação***

As demais tarefas que não são realizadas diretamente a partir da visualização de redes, são levantadas na identificação das tarefas de interação com a ferramenta. Para isso, é necessária a análise da ferramenta que será avaliada para entender como os usuários interagem com ela, identificação dos fluxos de uso mais comuns, as principais funcionalidades e as operações críticas que os usuários realizam. As tarefas podem variar em complexidade, desde tarefas simples e diretas até tarefas mais complexas que exigem múltiplos passos ou decisões. Tarefas exploratórias permitem aos participantes explorar e descobrir funcionalidades da ferramenta.

#### ***4.3.3 Definir o perfil e a quantidade de participantes***

A definição do perfil dos participantes é realizada com base nos objetivos da avaliação e nas características da ferramenta em estudo. Os avaliadores devem buscar participantes que representem o público-alvo do sistema que será avaliado. Por exemplo, se o objetivo for verificar como usuários novatos aprendem a realizar determinadas tarefas utilizando o sistema, o avaliador deve recrutar usuários inexperientes no uso do sistema e na realização das tarefas em questão (BARBOSA *et al.*, 2021). A definição pode incluir fatores como a área de atuação, a experiência prévia e habilidades relevantes.

A quantidade de participantes em uma avaliação é determinada por uma combinação de considerações de acordo com o contexto. Existe um grande debate sobre o número mínimo de participantes para que o teste de usabilidade seja completo e robusto ao encontrar os problemas de usabilidade e que está fora do escopo desta pesquisa (NIELSEN, 2000; BORSCI *et al.*, 2013). Dependendo da probabilidade de encontrar problemas de usabilidade na ferramenta, da diversidade do público-alvo de usuários e quão completo desejar-se que seja o teste, o número de participantes pode variar bastante. Caso seja necessário obter resultados estatisticamente significativos, a amostra de usuários deve ser suficientemente grande e representativa. Entretanto, o tempo e outros recursos necessários para a coleta e análise de dados pode inviabilizar uma

abordagem estatística (BARBOSA *et al.*, 2021). Além disso, se o domínio da aplicação for muito específico, pode ser que seja impraticável a obtenção de uma amostra muito grande de usuários.

#### **4.3.4 Preparar os materiais de apoio**

De acordo com a definição do método da avaliação, o avaliador deve preparar os materiais de apoio necessários, esse material costuma incluir:

- Projeto de pesquisa, incluindo Termo de Consentimento Livre e Esclarecido (TCLE), submetido e aprovado no Comitê de Ética em Pesquisa (CEP). No Brasil, o projeto precisa ser submetido por meio da Plataforma Brasil<sup>18</sup>;
- Cenários para orientar os participantes sobre as tarefas que serão realizadas;
- Questionários ou roteiro de entrevistas pré-teste para coletar informações sobre o participante, como: área de atuação, conhecimento sobre o domínio e características pessoais;
- Questionários ou roteiro de entrevistas pós-teste para coletar informações sobre a opinião, as dúvidas, dificuldades ou desconfortos decorrentes da experiência de uso;
- Roteiro com todos os dados necessários para orientar a condução da avaliação, como: descrição do objetivo da avaliação, descrição do perfil dos participantes, os equipamentos, a infraestrutura de software necessária, a listagem dos materiais de apoio necessários, as instruções para o ambiente de teste e as atividades iniciais, além dos roteiros das entrevistas, os cenários de uso e as tarefas.

Quando for elaborar os cenários de tarefas, o avaliador deve estar atento ao tempo necessário para os participantes realizarem as tarefas, segundo Barbosa *et al.* (2021) o tempo estimado para um participante realizar cada tarefa não deverá ultrapassar 20 minutos.

#### **4.3.5 Teste-piloto**

Para avaliar o planejamento e analisar se a avaliação planejada produzirá os dados adequadamente, é necessário realizar um teste-piloto. O teste-piloto também permite ao avaliador verificar quanto tempo será necessário para a realização da avaliação.

O teste-piloto deve ser conduzido como se fosse uma sessão normal de avaliação, com um participante que não tenha envolvimento com o planejamento da avaliação e que

<sup>18</sup> <https://plataformabrasil.saude.gov.br/login.jsf>

possua o perfil dos participantes (BARBOSA *et al.*, 2021). Se algum problema for detectado, o planejamento e os materiais de apoio devem ser corrigidos. A execução do teste piloto é um processo iterativo podendo ser realizado até que tudo esteja pronto para a realização da avaliação definitiva.

Os dados do teste-piloto devem ser descartados, pois podem estar contaminados com problemas que ocorreram ou não serem compatíveis ou válidos considerando o planejamento da avaliação (BARBOSA *et al.*, 2021).

#### **4.3.6 Recrutar os participantes**

Uma vez que o teste piloto tenha sido bem sucedido, o próximo passo é recrutar os participantes. Para recrutar os participantes com os perfis definidos, o avaliador pode utilizar questionários ou entrevistas curtas a fim de conferir se uma pessoa possui o perfil desejado (BARBOSA *et al.*, 2021). Ao convidar os participantes o avaliador deve esclarecer brevemente quais os objetivos da avaliação, como será realizada e quanto tempo será requerido do participante. Confirmados a disponibilidade e o interesse, deve ser realizado o agendamento formalmente.

### **4.4 Condução da avaliação**

A condução da avaliação deve ocorrer conforme o planejamento e utilizando os materiais preparados. Para assegurar a validade do estudo, é importante que todos os participantes recebam o mesmo material, as mesmas informações e orientações.

Inicialmente, devem ser dadas boas-vindas, explicar ao participante o objetivo do estudo, o procedimento da avaliação e as orientações sobre o TCLE. Após isso, deve ser fornecido ao participante o TCLE para leitura e assinatura. Somente após a assinatura do termo, a avaliação prossegue com a entrega do cenário de uso com as tarefas.

Após as instruções iniciais, deve-se iniciar o registro dos dados por meio da gravação de áudio ou vídeo, gravação da tela do participante e anotações.

Em seguida, o participante responde ao questionário ou entrevista pré-teste. Caso seja um questionário, o avaliador pode ler as perguntas para o participante e registrar as respostas fornecidas, como numa entrevista estruturada, isso pode ajudar o participante a se sentir mais à vontade durante a avaliação (BARBOSA *et al.*, 2021).

Após isso, é iniciada a sessão de observação de uso das tarefas solicitadas em forma

de cenário. O avaliador pode realizar a leitura dos cenários e a descrição das tarefas para o participante realizá-las. Enquanto o participante interage com o sistema, são realizadas a observação direta da interação e a anotação de qualquer acontecimento que possa ser relevante para a interpretação dos dados coletados e eventuais dúvidas a serem esclarecidas na entrevista pós-teste. O participante não deve ser interrompido nem questionado enquanto realiza as tarefas solicitadas.

Terminada a sessão de observação de uso, o avaliador deve conduzir a entrevista pós-teste. A entrevista pós-teste é uma oportunidade para coletar a opinião do participante sobre a experiência de uso que acabou de vivenciar e esclarecer eventuais dúvidas sobre seu comportamento e suas intenções durante a execução das tarefas (BARBOSA *et al.*, 2021). Nessa etapa, também podem ser coletados dados como: sugestões de novas funcionalidades e o interesse em utilizar a ferramenta no dia-a-dia.

#### **4.5 Análise dos dados**

Na análise de dados, o avaliador revisa, interpreta e analisa os dados coletados durante a condução da avaliação, para entender como os usuários interagiram com o sistema e identificar áreas de melhoria.

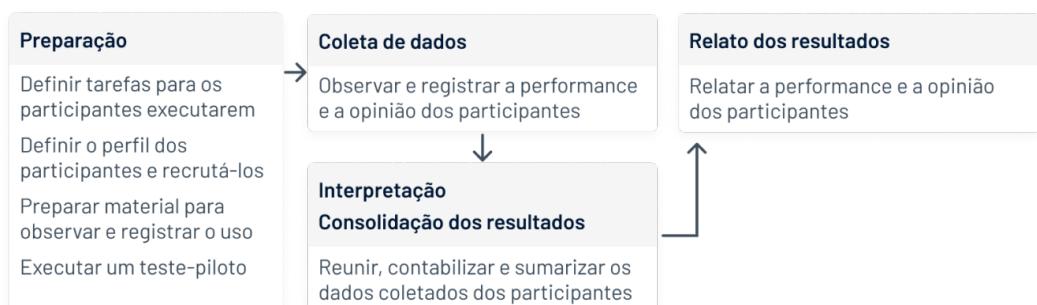
Os dados coletados incluem métricas quantitativas e dados qualitativos. Como métricas quantitativas, podem ser analisados o tempo de conclusão de tarefas, o grau de sucesso da execução das tarefas, a quantidade de erros, o número de vezes que a ajuda foi consultada, e assim por diante. Se questionários de usabilidade que utilizam escalas likert forem aplicados é possível computar um resultado quantitativo (escores). Alguns questionários fornecem métodos específicos para o cálculo de um escore final, mas normalmente também são usadas medidas de agregação como soma ou média. Os dados qualitativos são analisados por meio da transcrição das entrevistas e da observação de uso, os dados incluem: *feedbacks*, informações sobre a experiência do usuário, percepções, frustrações e sugestões.

Na análise de dados, também é possível identificar: problemas de interação e interface; necessidade da inclusão de contexto como: rótulos, descrições e legendas; necessidade de desenvolver recursos educacionais ou tutoriais para ajudar os usuários a entenderem como interpretar e interagir com as visualizações de dados; além de dificuldades recorrentes e comportamentos inesperados.

Figura 9 – Comparaçao entre a abordagem proposta e as atividades do teste de usabilidade definidas por Barbosa *et al.* (2021)



(a) Fluxo de execução da abordagem proposta para avaliação de ferramenta de visualização de rede



(b) Atividades do teste de usabilidade definidas por Barbosa *et al.* (2021)

Fonte: Elaborado pelo autora

## 4.6 Considerações Finais

Neste capítulo, foi apresentada a metodologia de avaliação proposta neste trabalho. A adaptação de testes de usabilidade para avaliar a visualização de redes multivariadas tem como principal diferencial em relação aos testes de usabilidade clássicos (Capítulo 2.3.2) a etapa de preparação ou planejamento (Figura 9). No planejamento da avaliação de redes multivariadas, há a inclusão da atividade de identificar as tarefas de análise da visualização, associando-as a uma taxonomia. Posteriormente, no Capítulo 6, a metodologia é usada para avaliar a aplicação PS-iMAP, uma aplicação real na área de patologia computacional. O PS-iMAP é descrito no capítulo a seguir.

## 5 PS-IMAP

O projeto PathoSpotter<sup>19</sup> é um projeto da Fiocruz da Bahia, iniciado em 2014, que reúne um grupo de patologistas e especialistas em ciência da computação com o objetivo de construir um conjunto de ferramentas para auxiliar patologistas em suas atividades diagnósticas e de pesquisa. Os desenvolvimentos mais significativos do PathoSpotter são na área de patologia renal, já que essa é a área de especialidade dos patologistas atuando no grupo e também graças à disponibilidade de uma robusta biblioteca de imagens digitais de lesões renais histológicas<sup>20</sup> (OLIVEIRA *et al.*, 2022). As imagens histológicas carregam uma variedade de informações sobre a doença do paciente, contendo informações relevantes para a definição do diagnóstico e, consequentemente, para o prognóstico da doença (OLIVEIRA *et al.*, 2022).

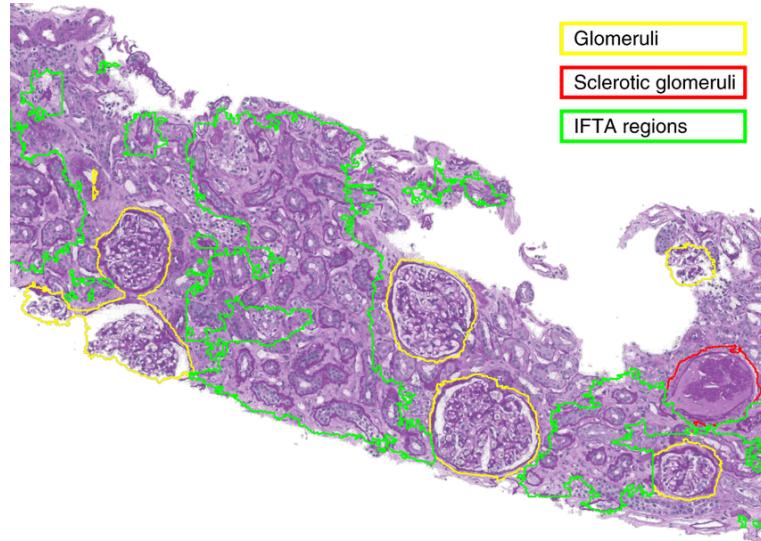
O projeto pretende oferecer ferramentas para facilitar o consenso de clínicos e patologistas, ajudando a obter diagnósticos mais precisos e rápidos; facilitar o estudo de correlações entre lesões e doenças; e auxiliar patologistas e pesquisadores de ciência da computação na classificação de lesões em biópsias renais.

A área em que o PathoSpotter se insere é chamada de Patologia Computacional, uma área de estudo que tem por finalidade integrar medicina e ciência da computação. Alguns dos problemas que os pesquisadores do PathoSpotter enfrentam são a dificuldade de obter informação do estado da arte da outra área (medicina e ciência da computação) e de comunicar seus próprios conhecimentos em uma equipe multidisciplinar. Um exemplo dessa dificuldade é o desenvolvimento de um software que consiga detectar Glomerulosclerose, uma condição de saúde que causa alterações morfológicas (incluindo esclerose) de pequenos vasos sanguíneos, conhecidos como glomérulos, nos rins (SANTOS *et al.*, 2024). Para isso, inicialmente patologistas precisam descrever o que é um glomérulo, como ele se manifesta em imagens obtidas a partir de lâminas de biópsias, quais os tipos de colorações das imagens, e depois explicar os tipos de alterações morfológicas que podem aparecer em casos de Glomerulosclerose. Por outro lado, patologistas também precisam entender como algoritmos de aprendizagem de máquina funcionam. Nesse caso, os algoritmos precisam antes segmentar (localizar) apenas os glomérulos em lâminas de biópsia (Figura 10) para serem mais eficientes. E muitas vezes uma lâmina de biópsia possui vários glomérulos apresentando alterações morfológicas distintas, como mostrado na Figura 10, além de possuir outros possíveis diagnósticos.

<sup>19</sup> <https://pathospotter.bahia.fiocruz.br/>

<sup>20</sup> A histologia estuda a estrutura microscópica e as funções das células, tecidos e órgãos.

Figura 10 – Imagem de lâmina de biópsia exibindo algumas estruturas histológicas já segmentadas: dentre as quais glomérulos saudáveis, escleróticos e fibrose intersticial e atrofia tubular (IFTA)



Fonte: Lutnick *et al.* (2019).

Neste contexto, ferramentas que auxiliem nessa troca de conhecimentos podem fomentar a comunicação e a colaboração entre esses pesquisadores. Assim, os pesquisadores do PathoSpotter decidiram desenvolver a aplicação PS-iMAP, que utiliza visualização de dados de associações de lesões e doenças para melhorar a comunicação e a colaboração entre patologistas e pesquisadores de ciência da computação em projetos na área de patologia computacional, fornecendo conhecimento fundamental nas duas áreas.

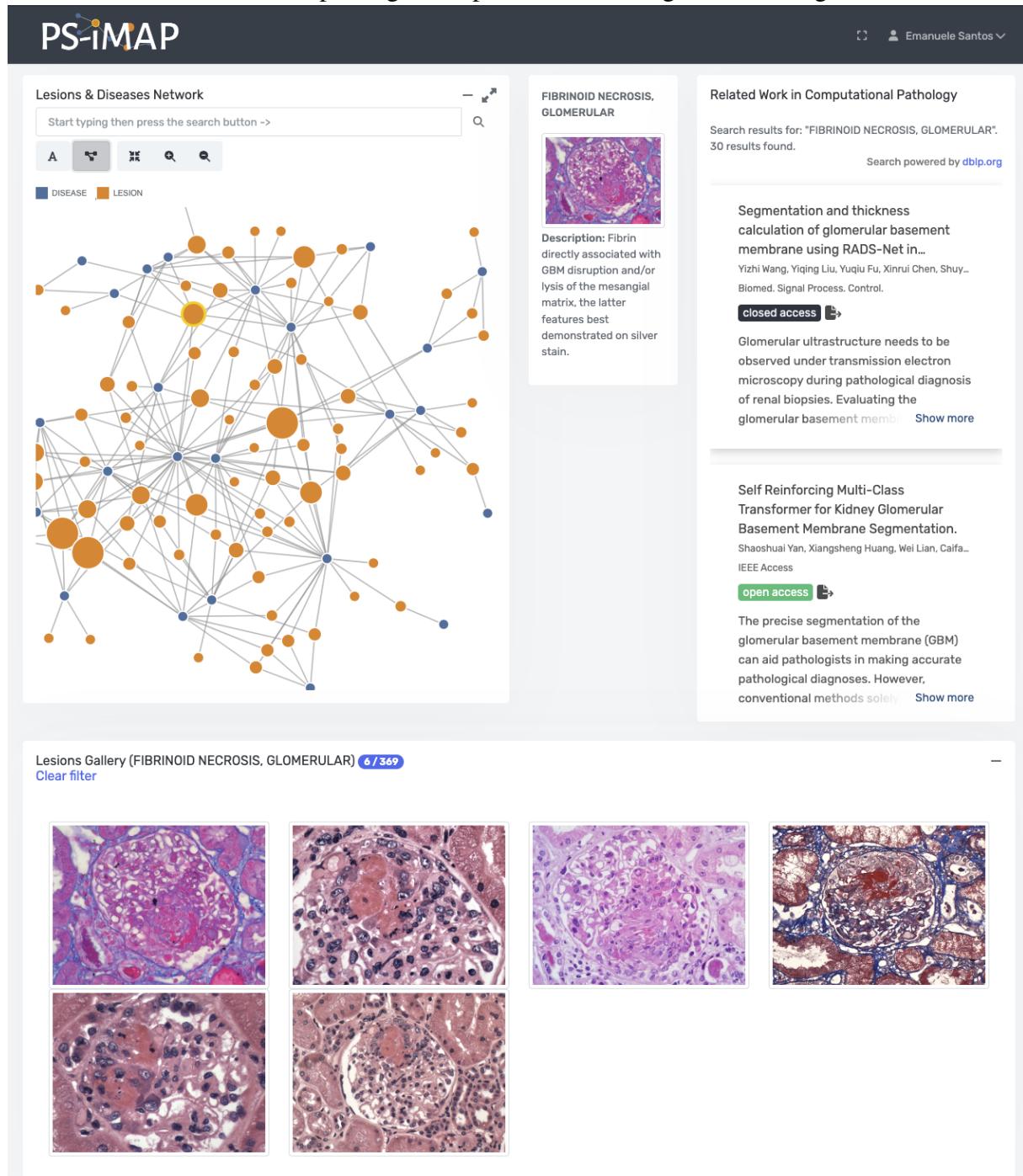
O PS-iMAP<sup>21</sup> é um sistema web, composto primariamente por dois componentes principais: um módulo *front-end* e um módulo *back-end*. O *front-end* possui três componentes visuais: uma rede multivariada de lesões e doenças representada em um diagrama nó-link, um painel para exibição de trabalhos relacionados de patologia computacional e uma galeria de imagens de lesões (Figura 11).

### 5.1 Rede multivariada de lesões e doenças

O primeiro componente visual, a rede multivariada de lesões e doenças, atualmente possui 78 lesões. A rede caracterizada como diagrama nó-link, representando a relação entre lesões e doenças renais, é considerada uma rede pequena e heterogênea (Figura 12). Na rede, o tipo de nó (lesão ou doença) é codificado por cor, as doenças são representadas em azul e as lesões em laranja. No caso de nós do tipo lesão, a quantidade de conexões é codificada pelo

<sup>21</sup> Um vídeo de demonstração está disponível em [https://www.youtube.com/watch?v=7R3frBOR\\_rI](https://www.youtube.com/watch?v=7R3frBOR_rI)

Figura 11 – Captura de tela do PS-iMAP mostrando seus três componentes visuais: uma rede multivariada de lesões e doenças representada em um diagrama nó-link, um painel para exibição de trabalhos relacionados de patologia computacional e uma galeria de imagens de lesões.



Fonte: Elaborado pela autora

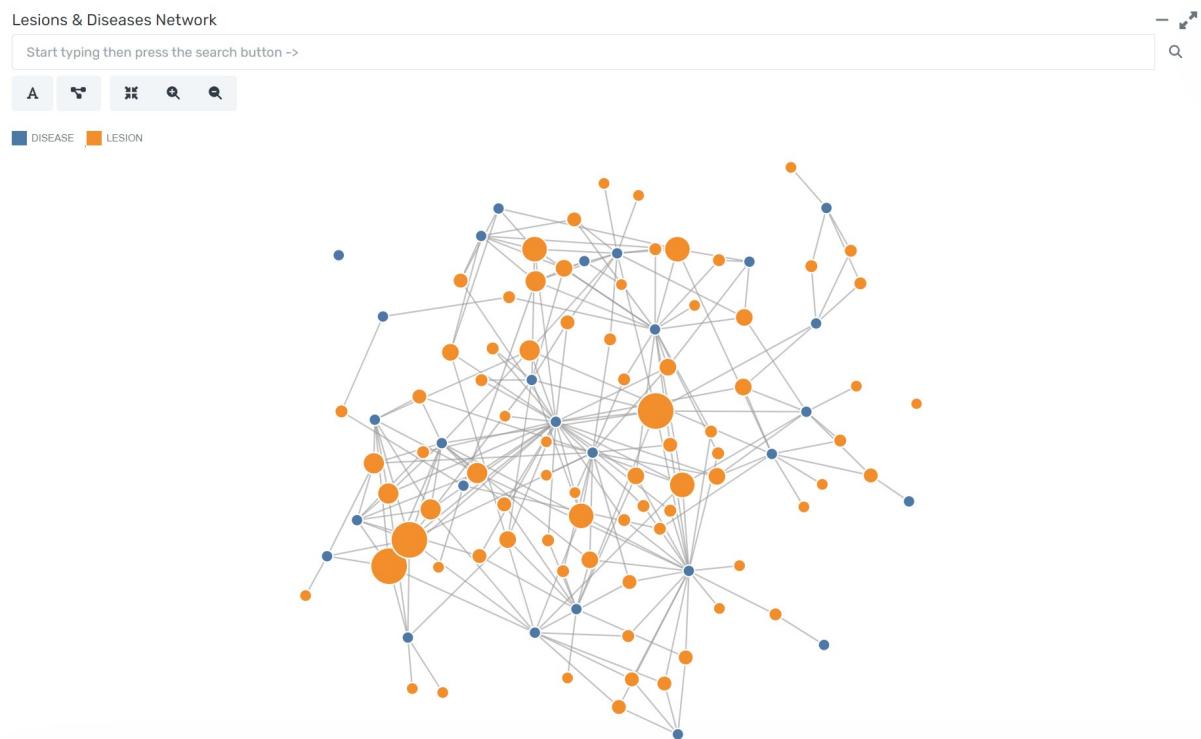
tamanho do nó, isso significa que nós maiores representam lesões mais prevalentes. Já os nós das doenças possuem tamanho constante. O layout da rede não é restrito, ou seja, é um layout livre, utilizando o posicionamento direcionado à força<sup>22</sup>.

<sup>22</sup> Os métodos direcionados à força posicionam os nós do gráfico de modo que todas as arestas tenham mais ou menos comprimento igual e há o menor número possível de cruzamentos de arestas, atribuindo forças como

Na rede, as lesões estão sempre ligadas a doenças. Não existe ligação entre lesões ou entre doenças. Dessa forma, pode-se observar como uma mesma lesão está ligada a diversas doenças distintas e como uma mesma doença pode estar associada a múltiplas lesões.

Além do diagrama nó-link, no componente de rede há um campo de busca por lesão ou doença e uma lista de funcionalidades disponíveis na barra de ferramentas no topo da rede. Estão disponíveis as funcionalidades de exibir os rótulos dos nós, o modo focado e navegação. No modo focado, para facilitar a observação de doenças ligadas a uma dada lesão, destacam-se apenas o nó selecionado e os seus vizinhos imediatos. Para ajudar na navegação, também estão disponíveis as funções para manipular a visualização por meio de *zoom in* e *out* e a possibilidade de restaurar a visualização para sua forma inicial (Figura 12).

Figura 12 – Captura de tela do componente visual da rede multivariada. No componente há um campo de busca por lesão ou doença, uma barra de ferramentas e um diagrama nó-link representando a relação entre lesões e doenças renais



Fonte: A autora.

Ao passar o mouse sobre um nó, é exibido um *tooltip* com o rótulo do nó, seu tipo e quantas conexões aquele nó possui. Ao selecionar um nó de lesão, é mostrado à direita da rede, quando disponível, um *card* de detalhes com uma imagem em miniatura e uma descrição da

---

se as arestas funcionassem como molas e os nós como partículas eletricamente carregadas. O gráfico inteiro é então simulado como um sistema físico (HOLTEN; WIJK, 2009).

lesão selecionada (Figura 13). A seleção de uma lesão também provoca a busca de trabalhos relacionados e a galeria é filtrada para exibir somente as imagens referentes àquela lesão.

Outra funcionalidade da rede é a ligação com atlas renais e outras fontes abertas sobre doenças renais. Nesse caso, para saber mais sobre uma determinada doença, um clique duplo no nó correspondente vai abrir uma página com as informações sobre aquela doença (Figura 14).

## 5.2 Painel de trabalhos relacionados

Nos trabalhos relacionados (Figura 15), é possível observar a quantidade de trabalhos encontrados e, para cada trabalho encontrado, é exibido: o título, os autores, onde o artigo foi publicado, um trecho do *abstract*, um indicador de se o trabalho completo está disponível livremente para leitura (*open access*) ou não (*closed access*), e a possibilidade de abrir o artigo em uma nova aba.

## 5.3 Galeria de imagens de lesões

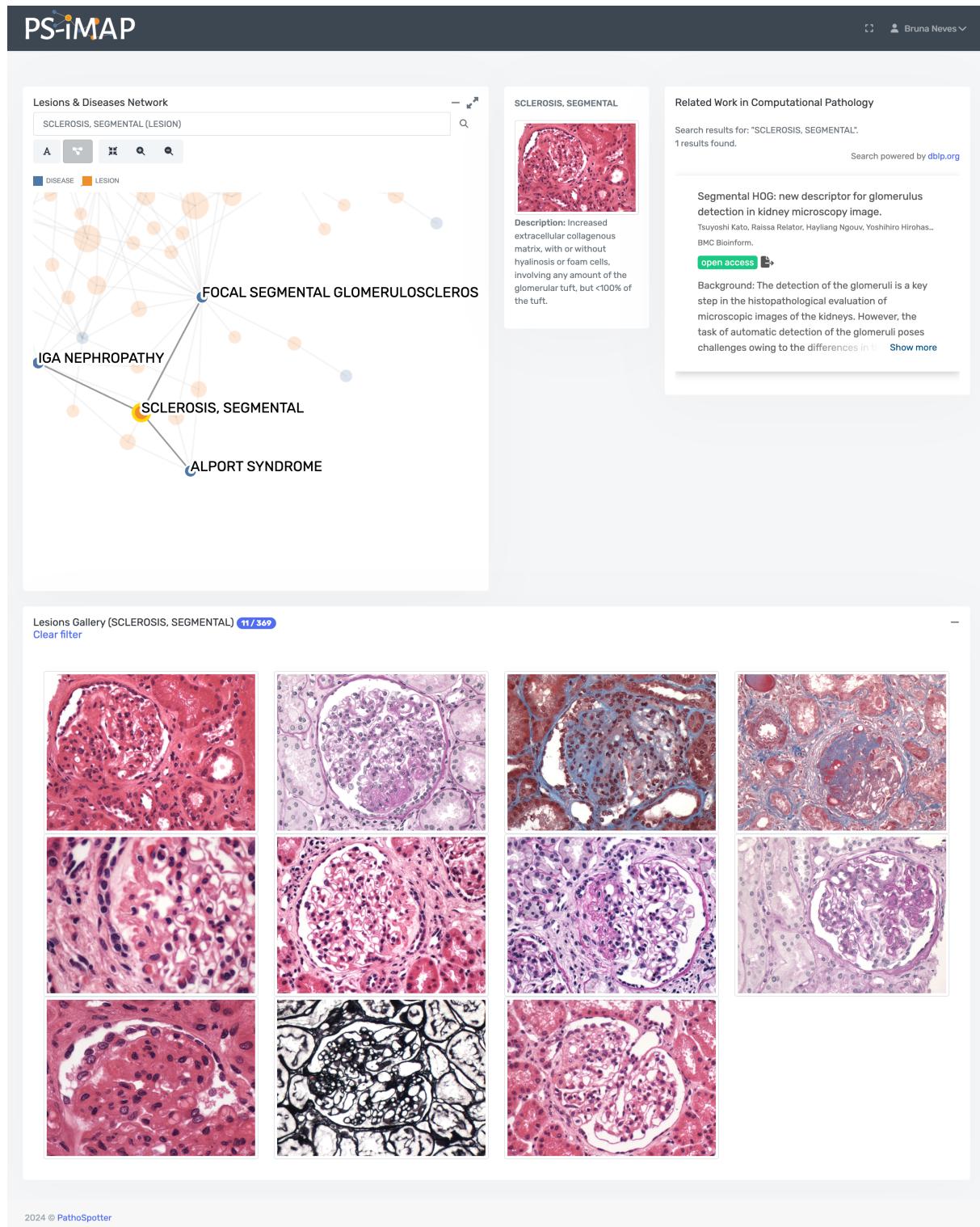
A galeria de imagens de lesões (Figura 16) atualmente possui 369 imagens, com diversas colorações e exemplos diferentes para um melhor entendimento de cada lesão. A galeria também está sincronizada com a rede e ao selecionar um nó de lesão na rede, a galeria é filtrada para exibir somente as imagens referentes àquela lesão. A galeria informa a quantidade total de imagens disponíveis e as funcionalidades de limpar o filtro e minimizar. Ao clicar em uma imagem da galeria, ela é mostrada em seu tamanho original.

## 5.4 Considerações Finais

Neste capítulo, foi apresentada o sistema web PS-iMAP. O PS-iMAP facilita a comunicação entre patologistas e cientistas da computação através de três componentes visuais: uma rede multivariada de lesões e doenças representada em um diagrama nó-link, um painel para exibição de trabalhos relacionados de patologia computacional e uma galeria de imagens de lesões.

Por meio do PS-iMAP, patologistas e cientistas da computação têm acesso facilitado às principais informações sobre lesões, suas manifestações clínicas e o estado da arte das técnicas de patologia computacional.

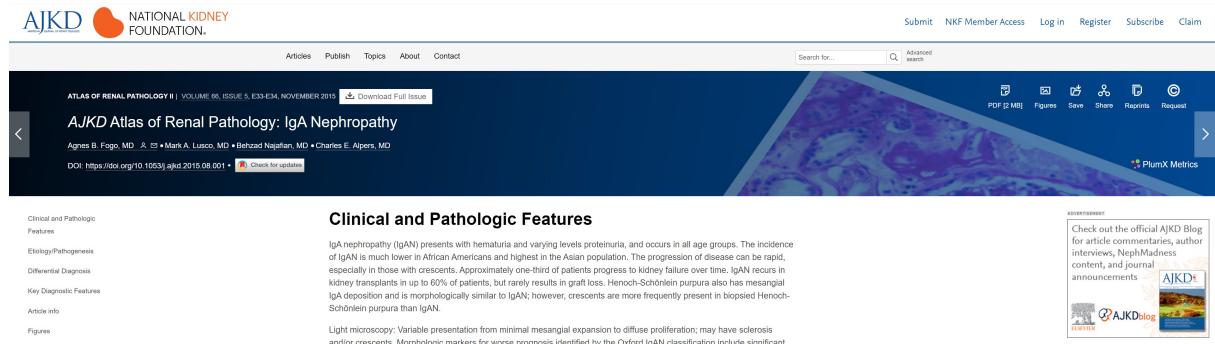
Figura 13 – Captura de tela do PS-iMAP após a seleção do nó *SCLEROSIS, SEGMENTAL* mostrando o *card* de detalhes com uma imagem em miniatura e uma descrição da lesão selecionada. No componente da rede, o modo focado está ativado, destacando-se apenas o nó selecionado e os seus vizinhos imediatos. No painel para exibição de trabalhos relacionados e na galeria de imagens estão sendo exibidos somente os resultados referentes à lesão *SCLEROSIS, SEGMENTAL*.



2024 © PathoSpotter

Fonte: A autora.

Figura 14 – Captura de tela do Atlas. A imagem mostra as características da Nefropatia por IgA.



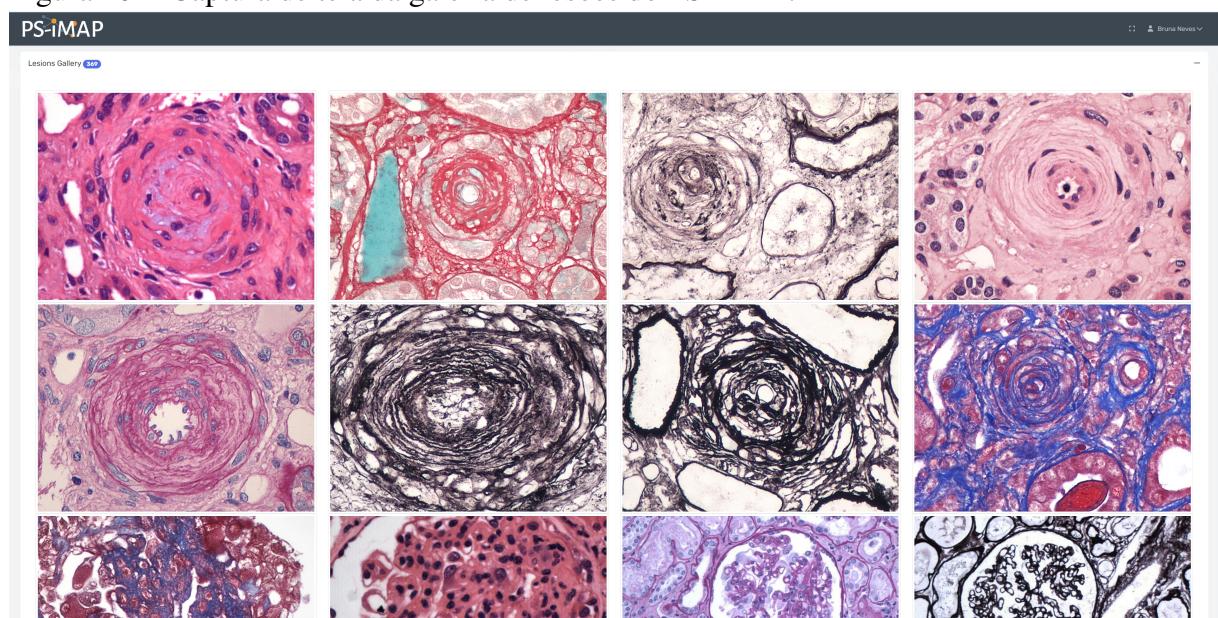
Fonte: A Autora a partir de American Journal of Kidney Diseases (2024).

Figura 15 – Captura de tela do componente de trabalhos relacionados em patologia computacional

Fonte: A autora.

No próximo capítulo, será descrita a avaliação do PS-iMAP, utilizando a metodologia de avaliação proposta.

Figura 16 – Captura de tela da galeria de lesões do PS-iMAP.



Fonte: A autora.

## 6 APLICAÇÃO DO TESTE DE USABILIDADE: ESTUDO DE CASO COM O PS-iMAP

Neste capítulo, a abordagem de avaliação com usuários, definida no Capítulo 4, será usada para avaliar a aplicação PS-iMAP. A Seção 6.1 mostra o procedimento metodológico da avaliação com os usuários. A Seção 6.2 mostra os resultados da avaliação.

### 6.1 Avaliação do PS-iMAP

Figura 17 – Fluxo de execução da avaliação realizada no PS-iMAP. No planejamento da avaliação, os participantes foram separados em dois grupos: com e sem treinamento.



Fonte: Elaborado pelo autora

O PS-iMAP foi avaliado utilizando a abordagem proposta neste trabalho (Capítulo 4) seguindo cinco etapas: definição dos objetivos, definição do método, planejamento, condução e análise dos dados (Figura 17).

A avaliação foi realizada durante o desenvolvimento da ferramenta, caracterizando-se como uma avaliação formativa, permitindo validar se as suposições feitas durante o design do produto realmente funcionam com os usuários. Durante a avaliação foram registrados dados sobre o desempenho dos participantes na realização das tarefas, suas opiniões decorrentes da experiência de uso do PS-iMAP e *insights* valiosos sobre necessidades e desejos que não foram previamente considerados, abrindo oportunidades para melhorias.

### **6.1.1 Definição dos objetivos da avaliação**

Primeiramente, na etapa de avaliação, foram definidos os seguintes objetivos:

- Identificar problemas na interação, na interface e na visualização que prejudiquem o uso, ou seja, a usabilidade.
- Verificar se a visualização de rede fornece as informações sobre doenças e lesões de forma comprehensível tanto para patologistas quanto para cientistas da computação.
- Verificar se os outros recursos interativos oferecem aos usuários o apoio computacional adequado.

A Tabela 5 detalha os objetivos por meio de questões mais específicas que a avaliação deverá responder.

**Tabela 5 – Perguntas que a avaliação buscou responder**

<b>Objetivo</b>	<b>Perguntas a serem respondidas</b>
Identificar problemas na interação, na interface e na visualização	O usuário consegue operar o sistema? Ele atinge o seu objetivo? Em quanto tempo? Após cometer quantos erros? Que parte da interface, da interação e da visualização o deixa insatisfeito? As funcionalidades interativas (zoom, filtros, etc.) são úteis e de fácil acesso?
Verificar se a visualização de rede fornece as informações sobre doenças e lesões de forma comprehensível tanto para patologistas quanto para cientistas da computação	As informações apresentadas são relevantes para o usuário? Os usuários comprehendem a informação apresentada? Os usuários conseguem realizar tarefas específicas de maneira eficiente usando a visualização? O quanto eles são motivados a explorar as funcionalidades da visualização?
Verificar se os outros recursos interativos oferecem aos usuários o apoio computacional adequado	De que maneira os usuários utilizam a ferramenta? Em que difere do planejado? O quanto os usuários consideram o apoio computacional adequado na realização de suas atividades?

Fonte: Elaborada pela autora.

### **6.1.2 Escolha da modalidade da avaliação**

A avaliação foi aplicada de maneira remota para atingir os potenciais usuários de diferentes localizações geográficas e de maneira mais econômica, eliminando a necessidade de deslocamento e de uma infraestrutura física específica (como laboratórios ou salas de testes); e moderada com o acompanhamento direto de um moderador para entender como os usuários interagem com as funcionalidades em tempo real, guiando as tarefas e fazendo perguntas quando necessário.

A avaliação com o usuário foi realizado em 5 fases: entrevista estruturada pré-teste para a coleta de dados sobre os participantes, treinamento sobre a ferramenta, exploração livre para coleta das primeira percepção dos participantes, observação de uso das tarefas de análise de rede e de interação com a interface e entrevista pós-teste para a coleta de opiniões, *feedbacks* e sugestões.

### **6.1.3 Planejamento**

Para uma condução adequada de uma avaliação que forneça resultados confiáveis e úteis, foi realizado o planejamento da avaliação, onde realizaram-se as atividades: descrição das tarefas, definição do perfil e da quantidade de participantes, preparação dos materiais de apoio, execução do teste piloto e recrutamento.

#### **6.1.3.1 Descrição das tarefas**

Para realizar a avaliação, os participantes foram convidados a realizar um conjunto de tarefas de análise de rede e um conjunto de tarefas de interação com a ferramenta. Para isso, houve várias discussões entre os patologistas e os cientistas da computação do PathoSpotter para definir as tarefas que seriam realizadas durante a avaliação. Nessa fase, a preocupação maior era em definir não somente as tarefas de domínio, mas também considerar a capacidade da representação da rede em atendê-las, associando-as com as tarefas da taxonomia de rede.

Após as discussões, as tarefas foram organizadas em três tipos, conforme a Tabela 6. O resultado foi uma lista de 13 tarefas a serem realizadas pelos participantes: 7 tarefas de análise de rede, 5 tarefas de interação com outros recursos disponíveis na ferramenta e 1 tarefa especial envolvendo tanto a rede como os outros recursos disponíveis. Antes de realizar as tarefas, o participante deve realizar o login na aplicação.

Todos os participantes realizaram as mesmas tarefas, com exceção da Tarefa 12 (T12) que é diferente, dependendo do perfil. A necessidade dessa diferenciação é porque o interesse no trabalho relacionado seria diferente: patologistas estariam mais interessados em que tarefas do seu dia-a-dia estariam sendo automatizadas e os cientistas da computação estariam interessados nos métodos e técnicas computacionais empregados no artigo.

A Tarefa Especial (TE) é fazer com que os participantes interajam com a ferramenta sem seguir um roteiro ou instruções específicas, fornecendo *insights* valiosos sobre como os participantes navegam, quais funcionalidades exploram primeiro, e onde encontram dificuldades ou surpresas.

Para cada tarefa realizada por cada participante, são coletados: sucesso (ou não) da sua execução e o tempo necessário para concluir-la. Além disso, também são registradas as anotações do avaliador e a captura de áudio e vídeo da interação completa.

Tabela 6 – Lista de instrução da tarefa e o seu tipo: análise de rede ou interação com a ferramenta

Instrução da tarefa	Tipo de tarefa
<b>TE</b> Explore a ferramenta livremente e relate suas descobertas (Há algo surpreendente ou particularmente interessante?)	Análise de rede e interação com a ferramenta
<b>T01</b> Localize a lesão <i>HYPERCELLULARITY, ENDOCAPILLARY</i>	Análise de rede
<b>T02</b> Encontre todas doenças ligadas à lesão <i>HYPERCELLULARITY, ENDOCAPILLARY</i>	Análise de rede
<b>T03</b> Selecione a lesão <i>SCLEROSIS, SEGMENTAL</i>	Análise de rede
<b>T04</b> Explore a rede de maneira focada	Interação com a ferramenta
<b>T05</b> Liste as doenças ligadas à lesão <i>SCLEROSIS, SEGMENTAL</i>	Análise de rede
<b>T06</b> Encontre a lesão com o maior número de doenças	Análise de rede
<b>T07</b> Localize a doença <i>IGA NEPHROPATHY</i>	Análise de rede
<b>T08</b> Descubra as características da doença <i>IGA NEPHROPATHY</i>	Interação com a ferramenta
<b>T09</b> Localize a lesão <i>SCLEROSIS, SEGMENTAL</i>	Análise de rede
<b>T10</b> Descubra quantas imagens estão associadas à lesão <i>SCLEROSIS, SEGMENTAL</i>	Interação com a ferramenta
<b>T11</b> Identifique quantos resultados de trabalhos relacionados à lesão <i>SCLEROSIS, SEGMENTAL</i> foram encontrados e indique os que são de acesso aberto	Interação com a ferramenta
<b>T12</b> Para os cientistas da computação: Qual(is) a(s) principal(is) técnica(s) computacional(is) utilizada(s) nos primeiros 3 trabalhos relacionados à lesão <i>SCLEROSIS, SEGMENTAL</i> ?  Para os patologistas: Quais as atividades do patologista que já foram automatizadas nos primeiros 3 trabalhos relacionados à lesão <i>SCLEROSIS, SEGMENTAL</i> ?	Interação com a ferramenta

Fonte: Elaborada pela autora.

#### 6.1.3.1.1 Tarefas de análise da rede

A definição das tarefas de análise de rede foi realizada com o entendimento das técnicas de avaliação e das taxonomias de tarefas existentes para o tipo de visualização disponível no PS-iMAP, atendendo as tarefas de domínio e as particularidades da visualização. Conforme recomendado pela metodologia proposta, as tarefas foram definidas com base na taxonomia de Nobre *et al.* (2019), já que se trata de uma taxonomia de tarefas recente e com foco específico em análise de rede multivariada. Dessa forma, para cada tarefa do domínio, foi realizada a associação com tarefas topológicas correspondentes. Por exemplo, a tarefa “T01 - Localize a lesão *hypercellularity, endocapillary*” é uma tarefa de identificação de um nó (estrutura topológica afetada) com o atributo nome = ‘*hypercellularity, endocapillary*’.

A Tabela 7 apresenta o resultado desse processo de associação para as 6 tarefas de análise de rede já listadas na Tabela 6, abrangendo as estruturas topológicas nós únicos e vizinhos e o tipo de tarefa, de acordo com a Tabela 1.

Tabela 7 – Lista das tarefas de análise de rede, estrutura topológica e tipo de tarefa de acordo com a taxonomia de Nobre *et al.* (2019)

Instrução da tarefa	Estrutura topológica (tipo de tarefa) (NO-BRE <i>et al.</i> , 2019)
<b>T01</b> Localize a lesão <i>HYPERCELLULARITY, ENDOCAPILLARY</i>	Nó (TgA)
<b>T02</b> Encontre todas doenças ligadas à lesão <i>HYPERCELLULARITY, ENDOCAPILLARY</i>	Vizinhos (AgT)
<b>T03</b> Selecione a lesão <i>SCLEROSIS, SEGMENTAL</i>	Nó (TgA)
<b>T05</b> Liste as doenças ligadas à lesão <i>SCLEROSIS, SEGMENTAL</i>	Vizinhos (AgT)
<b>T06</b> Encontre a lesão com o maior número de doenças	Nó (TgA)
<b>T07</b> Localize a doença <i>IGA NEPHROPATHY</i>	Nó (TgA)
<b>T09</b> Localize a lesão <i>SCLEROSIS, SEGMENTAL</i>	Nó (TgA)

Fonte: Elaborada pela autora.

#### 6.1.3.1.2 Tarefas de interação

Para a definição das tarefas de interação com a ferramenta, foi necessária uma análise da ferramenta e das suas principais funcionalidades, como apresentado no Capítulo 5. As tarefas variavam em complexidade, desde tarefas simples e diretas até tarefas mais complexas que exigiam múltiplos passos.

As tarefas de interação foram:

- T04 - Explore a rede de maneira focada

- T08 - Descubra as características da doença *IGA NEPHROPATHY*
- T10 - Descubra quantas imagens estão associadas à lesão *SCLEROSIS, SEGMENTAL*
- T11 - Identifique quantos resultados de trabalhos relacionados à lesão *SCLEROSIS, SEGMENTAL* foram encontrados e indique os que são de acesso aberto.
- T12 - Para os cientistas da computação: Qual(is) a(s) principal(is) técnica(s) computacional(is) utilizada(s) nos primeiros 3 trabalhos relacionados à lesão *SCLEROSIS, SEGMENTAL*? / Para os patologistas: Quais as atividades do patologista que já foram automatizadas nos primeiros 3 trabalhos relacionados à lesão *SCLEROSIS, SEGMENTAL*?

#### *6.1.3.2 Definição do perfil e da quantidade de participantes*

A definição do perfil dos participantes foi realizada com base nos objetivos da avaliação e nas características da ferramenta em estudo, representando o público-alvo do sistema que foi avaliado (BARBOSA *et al.*, 2021): patologistas e cientistas da computação com diferentes níveis de experiência.

Foram definidos 8 participantes: 4 patologistas e 4 cientistas da computação.

#### *6.1.3.3 Preparação dos materiais de apoio*

Antes de iniciar a coleta dos dados, foram preparados os materiais de apoio necessários:

- TCLE (Apêndice A) submetido e aprovado no CEP.
- Entrevista estruturada pré-teste (Apêndice B) para coletar informações sobre o participante, como: área de atuação, escolaridade, conhecimento sobre o domínio entre outros.
- Cenários de uso para contextualizar os participantes sobre as tarefas: um para patologistas (Apêndice C) e outro para cientistas da computação (Apêndice D).
- Roteiro da entrevista semiestruturada pós-teste (Apêndice E) para coletar informações sobre a opinião, as dúvidas, dificuldades ou desconfortos, sugestões e o interesse em utilizar a ferramenta.
- Roteiro com todos os dados necessários para orientar a condução da avaliação (Apêndice F), como: descrição do objetivo da avaliação, descrição do perfil

dos participantes, os equipamentos, a infraestrutura de software necessária, a listagem dos materiais de apoio necessários, as instruções para o ambiente de teste e as atividades iniciais.

- Vídeo de treinamento <sup>23</sup>.

O estudo foi aprovado pelo CEP do Instituto Gonçalo Moniz - FIOCRUZ (Anexo A).

Os cenários foram divididos em três partes: exploração livre, interação com a visualização de dados e interação com outros recursos na ferramenta PS-iMAP. Quando elaborados os cenários de tarefas, foi considerado o tempo necessário para os participantes realizarem as tarefas, visto que, o tempo estimado para um participante realizar cada tarefa não deve ultrapassar 20 minutos (BARBOSA *et al.*, 2021).

Foi preparado um vídeo para a etapa de treinamento, apresentando uma introdução sobre a ferramenta e o seu objetivo, de acordo com a necessidade observada na execução de um teste-piloto.

#### *6.1.3.4 Testes-piloto*

Foram realizados dois testes-piloto para avaliar o planejamento e analisar se a avaliação planejada produziria os dados necessários para responder os objetivos. Os testes foram conduzidos como uma sessão normal de avaliação, com participantes que não tinham envolvimento com o planejamento da avaliação e nenhum conhecimento sobre a ferramenta avaliada. Os testes-piloto também permitiram verificar quanto tempo seria necessário para a realização da avaliação.

O primeiro teste piloto foi realizado com uma cientista da computação com 5 anos de experiência. O teste durou 33 minutos e a participante apenas não conseguiu realizar uma tarefa. Durante o teste foi analisado que a participante teve muitas dúvidas a respeito da ferramenta e que seria necessário uma etapa de treinamento para uma introdução rápida sobre o objetivo e funcionamento da ferramenta. No roteiro do questionário pré-teste foram adicionadas perguntas sobre o nível de escolaridade e a área de atuação. Além disso, foram realizados ajustes no layout da ferramenta, de acordo com o *feedback* da participante: a funcionalidade de expandir o painel da rede para ocupar mais espaço da tela e melhorias na apresentação dos trabalhos relacionados.

Com o vídeo de treinamento produzido e realizados os ajustes necessários no planeja-

---

<sup>23</sup> [https://youtu.be/7R3frBOR\\_rI](https://youtu.be/7R3frBOR_rI)

mento, o segundo teste piloto foi realizado. O participante era da área de ciências da computação, com 6 anos de experiência e com um doutorado em andamento na área de *machine learning*. O teste durou 35 minutos e o participante conseguiu realizar todas as tarefas. Durante a entrevista pós-teste o participante citou que provavelmente sem o vídeo teria sido mais difícil a realização das tarefas e que a tarefa de explorar a ferramenta teria sido mais interessante sem o vídeo.

A partir da análise dos dados coletados do segundo piloto, os pesquisadores decidiram acrescentar mais um objetivo à avaliação: avaliar se os participantes eram capazes de usar a ferramenta sem treinamento prévio e conhecimentos sobre a ferramenta, ou seja a facilidade de aprendizado da ferramenta (outro fator de usabilidade). Os participantes da avaliação foram divididos em dois grupos: com treinamento e sem treinamento. O participantes do grupo de treinamento assistiram ao vídeo explicativo, com aproximadamente 6 minutos de duração, antes de realizarem a tarefa de exploração livre, enquanto que os participantes do grupo sem treinamento, realizariam a tarefa de exploração livre logo após a entrevista pré-teste e não tiveram acesso ao vídeo em nenhum momento.

#### ***6.1.3.5 Recrutamento dos participantes***

Após a realização do teste-piloto, foram recrutados 8 participantes como definido na Seção 6.1.3.2, mas apenas 7 realizaram o teste (por problemas de agenda, um dos participantes não realizou o teste). Ao convidar os participantes foi explicado brevemente quais os objetivos da avaliação, como seria realizada e quanto tempo seria requerido. Confirmado a disponibilidade e o interesse, foi realizado o agendamento da avaliação formalmente. No agendamento, foi disponibilizado um link para a chamada de vídeo.

#### ***6.1.4 Condução da avaliação***

A condução da avaliação ocorreu conforme o planejamento e utilizando os materiais de apoio preparados. Inicialmente, foram dadas boas-vindas e enviado ao participante um link para uma página com os dados da avaliação: uma descrição sobre o projeto, o objetivo da avaliação, um link para o TCLE, o link para o PS-iMAP, os cenários de uso e as tarefas. Após isso, foram dadas instruções sobre como seria realizada a avaliação e as orientações sobre o TCLE.

Todos os participantes receberam as mesmas informações e orientações. Todas as avaliações tiveram a gravação de vídeo e da tela do participante.

No início da avaliação, para conhecer o participante, foi realizada a entrevista estruturada pré-teste, onde foram registrados dados como área de atuação, a tempo de experiência na área de atuação, o conhecimento sobre estrutura histológica, o uso de sites e a classificação do nível de dificuldade na caracterização desses tipos de estruturas, a frequência com que este usa computadores e a internet.

Para os participantes que estavam no grupo com treinamento, após o questionário pré-teste, foi enviado um link para o vídeo introdutório sobre a ferramenta e as suas funcionalidades. Durante essa etapa, para ter certeza de que o vídeo foi visualizado, foi pedido que os participantes compartilhassem a tela. Após assistir ao vídeo, era perguntado aos participantes se eles tinham dúvidas. Já os participantes que estavam no grupo sem treinamento, após o questionário pré-teste, foi iniciado o cenário com as tarefas.

Foi lido para o participante o cenário de exploração livre e solicitado que ele realizasse login na ferramenta e explorasse a ferramenta livremente relatando suas descobertas. Para essa exploração, foi solicitado que o participante pensasse em voz alta (*thinking aloud*). Não houve limite de tempo para a realização da exploração, o participante era informado que poderia finalizar quando achasse necessário, mas quando percebido que o participante havia interagido com a maioria das funcionalidades, ele era solicitado a iniciar as outras tarefas.

Depois da etapa de exploração livre, houve a sessão de observação de uso das tarefas solicitadas em forma de cenário, conforme detalhado nos Apêndices C e D. Enquanto o participante interagia com o sistema, foi realizada a observação direta da interação, a anotação de qualquer acontecimento que possa ser relevante para a interpretação dos dados coletados e eventuais dúvidas a serem esclarecidas na entrevista pós-teste.

Terminada a sessão de observação de uso, foi realizada a entrevista pós-teste. A entrevista pós-teste é uma oportunidade para coletar a opinião do participante sobre a experiência de uso que acabou de vivenciar e esclarecer eventuais dúvidas sobre seu comportamento e suas intenções durante a execução das tarefas. Nessa etapa, foram registradas a opinião sobre o PS-iMAP, as tarefas difíceis e fáceis da avaliação, termos desconhecidos, as dúvidas, dificuldades ou desconfortos, sugestões de melhorias, sugestões de novas funcionalidades e o interesse em utilizar a ferramenta no dia-a-dia.

### **6.1.5 Análise dos dados**

Com a finalização das avaliações, os dados coletados foram transcritos, interpretados e analisados, sendo relatados para a melhoria da ferramenta.

Os dados coletados incluíam métricas quantitativas e dados qualitativos. Como métricas quantitativas foram medidos o tempo de conclusão de tarefas e o grau de sucesso da execução das tarefas. Como dados qualitativos foram analisados:

- Comentários dos participantes sobre as primeiras impressões sobre a ferramenta.
- Primeira ação (ou clique).
- A execução das tarefas.
- Respostas a perguntas abertas sobre a experiência do usuário, sugestões e opiniões.

## **6.2 Análise dos Resultados da Avaliação**

A avaliação consistia em realizar as tarefas descritas nos cenários, que foram divididos em três partes: exploração livre, interação com a visualização de dados e interação com outros recursos da ferramenta PS-iMAP. A avaliação foi realizada com sete participantes: quatro patologistas e três cientistas da computação.

A seguir são descritos em detalhes os resultados das etapas da avaliação: entrevista estruturada pré-teste, exploração livre, observação de uso e entrevista semiestruturada pós-teste. Ao final da seção, encontra-se uma discussão sobre os resultados obtidos.

### **6.2.1 Entrevista estruturada pré-teste**

Para conhecer o perfil dos participantes da pesquisa, no início da avaliação foi realizada a entrevista estruturada pré-teste (Apêndice B), onde foram registrados dados como o tempo de experiência na área de atuação, o conhecimento sobre estrutura histológica<sup>24</sup>, o uso de sites e a classificação do nível de dificuldade na caracterização desses tipos de estruturas. Questionar sobre o entendimento dessas estruturas foi relevante para descobrir se os participantes sabem diferenciar lesões de doenças. A Figura 18 resume o perfil dos participantes.

Os participantes possuíam diferentes níveis de experiência na área de estudo e todos relataram que utilizavam o computador e a internet diariamente. Entre os participantes, a média

---

<sup>24</sup> Uma estrutura histológica é a estrutura microscópica de células, tecidos e órgãos.

Figura 18 – Resumo do perfil dos participantes



Fonte: Elaborado pelo autora

de idade foi 32 anos. Do total de participantes, a avaliação foi realizada apenas com duas participantes do sexo feminino.

Dos participantes da área da ciência da computação, dois sabiam o que é uma estrutura histológica, mas informaram não utilizar sites para visualizar as estruturas. Os participantes da área de medicina citaram diferentes sites para visualizar estrutura histológica, o mais citado foi o Pathology Outlines<sup>25</sup>.

Foi pedido aos participantes de medicina para definirem, em uma escala de 1 a 3, sendo 1 “dificuldade mínima” e 3 “dificuldade máxima”, o quanto eles sentem dificuldade na caracterização de uma estrutura histológica, todos os participantes responderam o valor 2, a justificativa mais comum para a resposta foi a de não existir exatidão para a classificação de todas as estruturas.

Os participantes da avaliação foram divididos em dois grupos: com treinamento e sem treinamento. Os participantes do grupo de treinamento assistiram ao vídeo explicativo, antes da realização da etapa de exploração livre, enquanto que os participantes do grupo sem treinamento iniciaram a etapa de exploração livre logo após a entrevista estruturada pré-teste.

<sup>25</sup> <https://www.pathologyoutlines.com/>

### 6.2.2 Exploração livre

A etapa de exploração livre foi realizada no primeiro cenário com as tarefas de realizar login e a Tarefa Especial (TE), instruindo aos participantes a explorar livremente a ferramenta e relatar qualquer descoberta que obtivessem, e que eles pensassem em voz alta (*thinking aloud*). O objetivo foi observar como os usuários explorariam e utilizariam o PS-iMAP de maneira natural e espontânea, sem pressão de tempo ou resultados, revelando comportamentos, interesses e dificuldades.

Com a verbalização dos seus pensamentos enquanto navegavam livremente, foi realizada uma análise de conteúdo (Apêndice G), com o objetivo de entender a primeira percepção dos participantes ao interagir, pela primeira vez, com a ferramenta. As unidades de análise foram 45 comentários feitos pelos participantes que foram categorizados em comentários de descoberta, sugestões, opiniões e dúvidas. Por meio dessa categorização, pode-se notar que a maioria dos comentários foram opiniões (22) e sugestões (12). Além disso, foram 10 comentários de descoberta e 5 dúvidas.

A Tabela 8 sumariza o tempo de duração da tarefa de exploração livre, a quantidade total de comentários e a categoria do primeiro comentário verbalizado durante a tarefa por participante.

Tabela 8 – Tempo da duração da tarefa de exploração livre, a quantidade total de comentários e a categoria do primeiro comentário realizado durante a tarefa por participante.

	P1	P2	P3	P4	P5	P6	P7
Tempo	10:27	04:50	06:35	06:45	04:15	02:07	07:19
Total de comentários	13	5	5	5	8	1	13
Categoria do primeiro comentário	Opinião	Descoberta	Opinião	Descoberta	Sugestão	Opinião	Descoberta

Fonte: Elaborada pela autora.

Além da primeira percepção do participante ao interagir com a ferramenta, também foi analisada a primeira ação (ou clique) de cada participante (Tabela 9). Todos os participantes no primeiro acesso à ferramenta, interagiram com a visualização de dados.

No vídeo de treinamento, após apresentar os três componentes da ferramenta (a rede multivariada, o painel de trabalhos relacionados de patologia computacional e a galeria de imagens de lesões), são mostrados mais detalhes sobre a rede multivariada. Os participantes que estavam no grupo com treinamento (P1, P2, P4 e P5) na primeira interação com a ferramenta, além de explorarem a rede (com a funcionalidade de *zoom* ou com um clique em uma lesão), navegaram rapidamente para outros recursos da ferramenta (*card* de detalhes da lesão ou a galeria

Tabela 9 – Primeira ação ou clique que o participante realizou ao interagir com o PS-iMAP

<b>Primeira interação</b>
<b>P1</b> <i>Zoom</i> na rede seguido de <i>scroll</i> para a galeria de imagens
<b>P2</b> Clicou em um nó de lesão e direcionou o mouse para os detalhes sobre ela
<b>P3</b> <i>Zoom</i> na rede
<b>P4</b> Clicou em um nó de lesão e direcionou o mouse para os detalhes sobre ela
<b>P5</b> <i>Zoom</i> na rede seguido de <i>scroll</i> para a galeria de imagens
<b>P6</b> Arrastou um nó de lesão
<b>P7</b> Clicou em um nó de lesão

Fonte: Elaborada pela autora.

de imagens), enquanto que os participantes sem treinamento manipularam por mais tempo a rede.

Durante a exploração livre, o participante 3 afirmou que a visualização de dados não estava clara para ele. O participante 4 ficou avaliando os conteúdos disponíveis no painel de trabalhos relacionados e nas imagens da galeria (verificando se existia ou não a lesão nas imagens). A participante 6 não deu *scroll* na página, não percebendo a galeria de imagens de lesões.

Um ponto bastante relevante foi que, durante a TE, todos os participantes tiveram opiniões positivas e a percepção de que eles ficaram envolvidos e motivados com a ferramenta procurando realmente utilizá-la e entendê-la. Ademais, por meio da análise da primeira ação ou clique, pode-se notar que o que mais chamou atenção no PS-iMAP foi a rede multivariada de lesões e doenças.

#### **6.2.3 Observação de uso**

Durante a observação de uso foram medidos o tempo de conclusão e o grau de sucesso da execução das tarefas (Tabela 10). A seguir, será apresentada a análise da execução de cada tarefa.

##### **T01 - Localize a lesão HYPERCELLULARITY, ENDOCAPILLARY**

Todos os participantes conseguiram realizar a tarefa com sucesso. Durante a observação, notou-se que alguns participantes esperavam que a ferramenta buscasse a lesão logo após selecioná-la ou digitá-la na barra de pesquisa. Um participante citou a funcionalidade do *autocomplete*.

P7 não utilizou a busca para encontrar a lesão, ela realizou a tarefa percorrendo as

Tabela 10 – Tempo da conclusão em minutos e o grau de sucesso da execução das tarefas. As tarefas executadas com sucesso são representadas em verde e as não sucedidas em vermelho

Instrução da tarefa	P1	P2	P3	P4	P5	P6	P7
<b>T01</b> Localize a lesão <i>HYPERCELLULARITY, ENDOCAPILLARY</i>	00:50	00:10	00:20	00:11	00:20	00:30	00:36
<b>T02</b> Encontre todas doenças ligadas à lesão <i>HYPERCELLULARITY, ENDOCAPILLARY</i>	00:38	00:31	01:24	00:25	00:30	01:18	00:22
<b>T03</b> Selecione a lesão <i>SCLEROSIS, SEGMENTAL</i>	00:12	00:27	00:32	00:19	00:11	00:17	02:09
<b>T04</b> Explore a rede de maneira focada	-	-	00:17	-	-	00:19	00:15
<b>T05</b> Liste as doenças ligadas à lesão <i>SCLEROSIS, SEGMENTAL</i>	00:10	00:09	00:06	00:41	00:09	00:10	00:09
<b>T06</b> Encontre a lesão com o maior número de doenças	01:09	00:21	00:46	01:05	00:18	00:49	02:38
<b>T07</b> Localize a doença <i>IGA NEPHROPATHY</i>	00:12	00:08	00:22	00:08	00:05	00:13	00:23
<b>T08</b> Descubra as características da doença <i>IGA NEPHROPATHY</i>	02:15	00:24	01:30	02:20	1:06	00:45	00:40
<b>T09</b> Localize a lesão <i>SCLEROSIS, SEGMENTAL</i>	00:15	00:10	00:10	00:04	00:10	00:19	00:37
<b>T10</b> Descubra quantas imagens estão associadas à lesão <i>IGA NEPHROPATHY</i>	00:05	00:07	00:02	00:02	00:02	00:20	00:35
<b>T11</b> Identifique quantos resultados de trabalhos relacionados à lesão <i>SCLEROSIS, SEGMENTAL</i> foram encontrados e indique os que são de acesso aberto	00:06	00:06	00:14	00:09	00:10	00:10	00:07
<b>T12</b> Para os cientistas da computação: Qual(is) a(s) principal(is) técnica(s) computacional(is) utilizada(s) nos primeiros 3 trabalhos relacionados à lesão <i>SCLEROSIS, SEGMENTAL</i> ? Para os patologistas: Quais as atividades do patologista que já foram automatizadas nos primeiros 3 trabalhos relacionados à lesão <i>SCLEROSIS, SEGMENTAL</i> ?	01:27	00:24	01:15	00:43	00:37	02:16	03:00

Fonte: Elaborada pela autora.

lesões da rede com o *mouse* até localizar a lesão correta.

#### ***T02 - Encontre todas doenças ligadas a lesão HYPERCELLULARITY, ENDOCAPILLARY***

Dois participantes (P3 e P7) que não tiveram a etapa de treinamento tiveram dificuldade e não conseguiram realizar T02. O participante P3 afirmou que “Olhando o gráfico, impossível” e que as doenças também poderiam ser disponibilizadas em uma tabela ao lado da visualização. P7 não citou todas as doenças (5) ligadas a lesão.

Todos participantes que tiveram a etapa de treinamento, utilizaram o modo focado, o que ajudou na execução da tarefa de maneira mais rápida. Já os participantes do grupo sem treinamento realizaram a tarefa em mais de 1 minuto, o tempo maior de execução dos participantes com treinamento foi de 38 segundos.

#### ***T03 - Selecione a lesão SCLEROSIS, SEGMENTAL***

Todos os participantes conseguiram realizar a tarefa com sucesso.

Novamente, P7 não utilizou a busca para encontrar a lesão, o que levou um tempo

maior para completar a tarefa, visto que ela teve que olhar para uma grande quantidade de nós para localizar a lesão correta.

#### ***T04 - Explore a rede de maneira focada***

Todos participantes que tiveram a etapa de treinamento (P1, P2, P3 e P4), utilizaram o modo focado desde o início das tarefas, por isso não foram solicitados a ativar a função.

Todos os participantes sem treinamento conseguiram habilitar o modo focado. Notou-se que os participantes que não tiveram acesso a etapa de treinamento demoraram para entender o que seria a “maneira focada” e alguns receberam o auxílio de que se tratava de uma funcionalidade na barra de ferramentas no topo da rede.

#### ***T05 - Liste as doenças ligadas à lesão SCLEROSIS, SEGMENTAL***

Com o modo focado ativado, todos os participantes conseguiram listar os nomes das doenças. Um participante patologista (P4) durante a tarefa analisou as doenças e apontou que estava faltando a doença Lúpus.

Possivelmente se os participantes P3 e P7, que não conseguiram realizar T02 e estavam no grupo sem a etapa de treinamento, soubessem sobre a funcionalidade, eles também teriam conseguido realizá-la.

#### ***T06 - Encontre as lesões com o maior número de doenças***

Um participante (P7) teve dificuldade de realizar T06, citando apenas uma das três lesões com maior número de conexões.

Para a realização da tarefa, o participante P1 precisou restaurar a visualização para sua forma inicial, o que custou um tempo maior para completar a tarefa.

Além disso, notou-se que, inicialmente, P4 respondeu baseado em seus conhecimentos e não na visualização afirmando que a lesão com maior número de doenças deveria ser *Interstitial Fibrosis*, mas logo depois ele se deu conta de que as lesões associadas ao maior número de doenças possuíam os maiores nós. P6 ficou com dúvida se a resposta era realmente de acordo com o tamanho do nó, mas ainda respondeu corretamente.

### ***T07 - Localize a doença IGA NEPHROPATHY***

Todos os participantes conseguiram localizar a doença com sucesso.

### ***T08 - Descubra as características da doença IGA NEPHROPATHY***

Nenhum participante conseguiu realizar T08, por mais que houvesse a explicação de como realizá-la na etapa de treinamento e que o segundo teste-piloto tenha conseguido realizá-la. Nessa tarefa, esperava-se que os participantes dessem um clique duplo no nó da doença para abrir uma página com as informações sobre ela.

Todos citaram que para descobrir as características da doença teriam que olhar todas as lesões ligadas a ela e os artigos relacionados. Um exemplo de resposta foi: “Eu iria primeiro ver as lesões ligadas a nefropatia por IGA, notaria que todas aqui tem [...]. Próximo passo seria olhar aqui [painel de trabalhos relacionados] os artigos sobre ela, porque não tem nada escrito aqui [*card* de detalhes], sobre a doença” (P5).

Para os participantes que tiveram acesso à etapa de treinamento, durante a entrevista pós-teste, foi comentado como seria a resolução da tarefa, os participantes citaram que não recordaram da funcionalidade e que seria mais intuitivo se a descrição da doença e a função para abrir uma página com mais informações sobre ela estivesse disponível no *card* de detalhes.

### ***T09 - Localize a lesão SCLEROSIS, SEGMENTAL***

Todos os participantes conseguiram realizar a tarefa com sucesso.

### ***T10 - Descubra quantas imagens estão associadas à lesão SCLEROSIS, SEGMENTAL***

P6 foi a única participante que não respondeu corretamente, já que não havia explorado ainda a galeria de imagens por não ter rolado a página. Ela respondeu que havia apenas uma imagem, a imagem disponível no *card* de detalhes sobre a lesão.

### ***11 - Identifique quantos resultados de trabalhos relacionados à lesão SCLEROSIS, SEGMENTAL foram encontrados e indique os que são de acesso aberto.***

Todos os participantes conseguiram realizar a tarefa com sucesso.

**T12 - Qual(is) a(s) principal(is) técnica(s) computacional(is) utilizada(s) nos primeiros 3 trabalhos relacionados à lesão SCLEROSIS, SEGMENTAL? / Quais as atividades do patologista que já foram automatizadas nos primeiros 3 trabalhos relacionados à lesão SCLEROSIS, SEGMENTAL?**

P7 foi a única participante que não conseguiu realizar essa tarefa. Notou-se que ela teve dificuldade de entender a pergunta, questionando algumas vezes sobre ela. Em algum ponto durante a tarefa, a participante parou de compartilhar a tela, há a hipótese de que ela fez isso para consultar um tradutor.

#### **6.2.4 Entrevista semiestruturada pós-teste**

Ao final da avaliação, foi realizada a entrevista pós-teste, onde foram registrados a opinião sobre o PS-iMAP, as tarefas difíceis e fáceis da avaliação, termos desconhecidos, as dúvidas, dificuldades ou desconfortos, sugestões de melhorias, sugestões de novas funcionalidades e o interesse em utilizar a ferramenta no dia-a-dia.

No início da entrevista pós-teste os participantes puderam expor sua opinião quanto à ferramenta. Um ponto bastante positivo foi a aceitação da ferramenta. Todos os participantes elogiaram a ferramenta, comentando de forma positiva sobre a visualização de dados ou as funcionalidades. Logo abaixo são apresentadas as respostas dos participantes quando perguntados a opinião sobre o PS-iMAP:

- “Eu acho que o objetivo está sendo cumprido de estreitar a área da computação com a área médica da patologia renal. Acho bem intuitiva essa análise via grafo, essa “teia de aranha” que você percebe as lesões e as doenças.” (P1).
- “Eu gostei, achei bem legal. Acho até que me ajudaria no doutorado também para ter uma visão mais macro sobre as lesões. É muito legal essa parte de artigos também. Achei legal mostrar imagens também de diferentes colorações. Achei bem bacana mesmo a ferramenta.” (P2).
- “Sobre o objetivo dela, eu achei bem legal [...] para o patologista deve ser uma baita ferramenta, enxergar a relação entre as doenças e as lesões, isso é bem legal” (P3).
- “Eu gostei, não é perfeita mas está no bom caminho.” (P4).
- “A ferramenta tá muito *user friendly* né? Tá muito boa. Tá massa, eu entendi o

objetivo dela que é para acessar as informações de uma forma mais fácil, de uma forma resumida, mais simples e mais fácil, principalmente para a galera que não é da área de patologia, né? para buscar as informações de maneira mais fácil. [...] É interessante pra gente olhar [...] ter fácil a exemplificação da lesão que a gente tá falando, porque, às vezes, você tá tendo uma discussão, até mesmo em uma chamada, em uma reunião online, que é muito difícil você explicar para pessoas que não veem isso todo dia, o que é determinada lesão e às vezes você ter isso aqui fácil para abrir e já exemplificar [...] acho que isso facilitaria muitos momentos que a gente travou assim durante reuniões. [...] Eu gostei bastante” (P5).

- “Achei ótima” (P6).
- “Achei bem interessante, achei útil, prática, assim, o fato de você poder pesquisar e chegar na lesão e ver as imagens e os artigos relacionados, eu acho que é bem interessante porque facilita muito a vida do patologista na rotina diária.” (P7).

É interessante notar que dos participantes que tiveram acesso à etapa de treinamento, três citaram o objetivo da ferramenta, quando perguntados a opinião.

Três participantes afirmaram que a tarefa mais difícil foi descobrir as características da doença *IGA NEPHROPATHY* (T08), dois participantes relataram que a tarefa mais difícil foi a décima segunda (T12) e os outros dois participantes afirmaram não ter tarefa difícil.

Sobre a tarefa mais fácil, alguns participantes citaram mais de uma tarefa. A maioria dos participantes (6) citou localizar lesão ou doença; T06 (encontrar as lesões com o maior número de doenças) foi citada duas vezes; T10 (descobrir quantas imagens estão associadas à lesão *SCLEROSIS, SEGMENTAL*) e T11 (identificar quantos resultados de trabalhos relacionados à lesão *SCLEROSIS, SEGMENTAL* foram encontrados e indique os que são de acesso aberto) foram citadas uma única vez.

Quando questionados sobre ter alguma dúvida sobre a ferramenta, todos os participantes responderam que não tinham dúvidas. Sobre as suas dificuldades ou desconfortos, os participantes falaram sobre a navegação por dois cliques para ver as informações da doença (apenas participantes do grupo com treinamento), a falta da descrição da doença de modo semelhante ao da lesão, a necessidade de precisar clicar na lupa ou na tecla enter no campo de busca, a falta de uma descrição mais detalhadas das imagens (coloração, *zoom* e lente) e a sobreposição de arestas na visualização.

A maioria das dificuldades ou desconfortos dos participantes pode ser resolvida implementando suas sugestões de melhoria. Os participantes sugeriram alterações ou recursos novos que poderiam ser feitos futuramente para melhorar a ferramenta. Estas sugestões estão listadas na Tabela 11.

**Tabela 11 – Resumo das sugestões de melhoria que os participantes apontaram para a ferramenta**

<b>Participantes</b>	<b>Sugestão de melhoria</b>
P1, P6, P4, P5, P6	Descrição mais detalhadas das imagens (coloração, zoom e lente)
P1, P2, P3, P4, P5	Descrição da doença
P3, P5	Diminuir a quantidade de imagem da galeria mostradas no primeiro acesso
P4, P7	Imagen de imunofluorescência
P2, P5	Pesquisa após selecionar a doença ou lesão
P4	Tabela/Lista com o nome das doenças associadas a lesão
P1	Adicionar o índice de qualidade nos artigos
P6	Acesso à imagem escaneada

Fonte: Elaborada pela autora.

Além das sugestões de melhoria, os participantes também sugeriram novas funcionalidades (Tabela 12).

**Tabela 12 – Resumo das sugestões de novas funcionalidades**

<b>Participantes</b>	<b>Sugestão de funcionalidade</b>
P2, P4	Pesquisar por duas lesões ou doenças
P1	Histórico de busca
P1	Anotação sobre a imagem
P5	Fórum de discussão nas imagens

Fonte: Elaborada pela autora.

Os participantes afirmaram que o PS-iMAP poderia ajudá-los em suas pesquisas e no seu trabalho. Notou-se o interesse dos participantes em utilizar a ferramenta:

- “Para mim que hoje trabalho mais com a parte de esclerose segmental e global, eu fiquei curioso para saber as doenças que estão mais associadas a essas lesões (...) saber uma descrição melhor da lesão, ver os trabalhos relacionados” (P1).
- “Acho que é uma ferramenta bastante informativa, né? Principalmente se for considerar que eu trabalho com umas quatro ou cinco lesões... daria para ter um resumão delas aqui [card de detalhes] e acho que até para extrapolar para por exemplo, sei lá, se eu pesquisar aqui [HYPERCELLULARITY, ENDOCAPIL-LARY] eu consigo ver que tem doenças associadas a ela e poderia dar alguma discussão interessante no artigo, ver a relação” (P2).
- “Ela seria ótima no meu dia-a-dia” (P3).

- “Durante a investigação de ‘qual é o nome daquela doença’, eu achei aquela alteração histológica, mas quais são as doenças que têm essa alteração histológica, às vezes é bom fazer uma listagem disso” (P4).
- “Quando tiver em reuniões, é uma ferramenta muito útil para deixar ela aberta ali, projetar na reunião e discutir ou até mesmo quando a gente tá nas reuniões presenciais mesmo para relembrar algumas coisas, revisar conteúdo é bom também. Ter um acesso rápido aos artigos referências sobre aquela doença, né?” (P5).
- “Com certeza essa ferramenta me ajudaria” (P6).
- “Essa ferramenta me ajudaria muito” (P7).

A equipe de pesquisadores do PathoSpotter foi surpreendida positivamente com a extensão e a qualidade dos *insights*, o engajamento dos participantes e as sugestões de ambos os tipos de participantes para a melhoria da ferramenta.

#### ***6.2.5 Resultados Quantitativos***

O tempo de execução das tarefas bem sucedidas estão na Figura 19. São mostrados os gráficos de pontos e os intervalos de confiança 95% por condição de treinamento (Figura 19-a e Figura 19-b) e por perfil dos participantes (Figura 19-c e Figura 19-d).

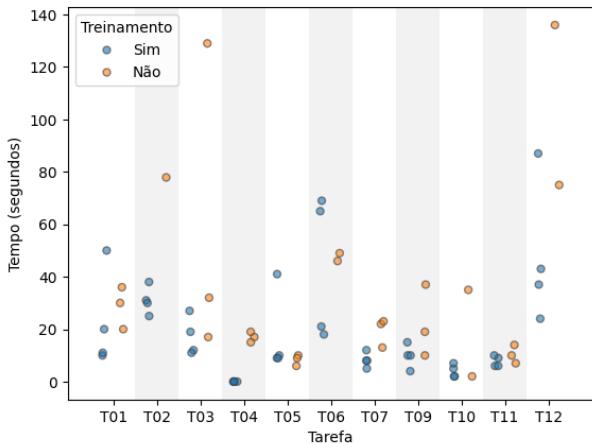
#### ***6.2.6 Discussão***

Após a descrição detalhada dos resultados produzidos, pode-se resumir em conjunto os fatos observados.

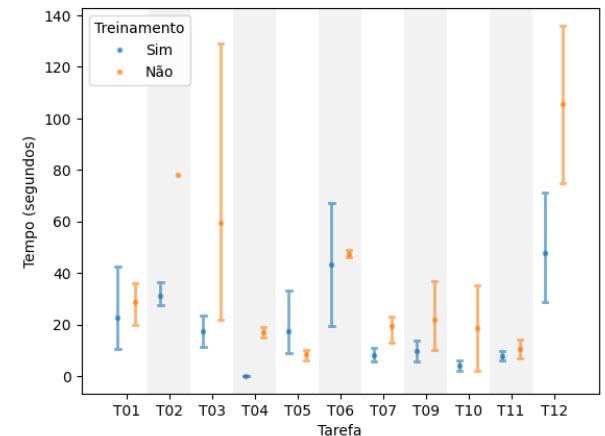
A respeito da etapa de treinamento, ela impactou no entendimento da ferramenta. Houve mais erros na execução das tarefas pelos participantes que não tiveram acesso ao treinamento. Além disso, os participantes com treinamento utilizaram, desde o início da avaliação, ferramentas que os auxiliaram na resolução das tarefas. No entanto, os participantes sem treinamento foram capazes de realizar a maioria das tarefas sem conhecimentos prévios indicando a facilidade de aprendizado e uso.

Sobre a execução das tarefas, não houve diferença significativa no tempo necessário para completá-las entre grupos de participantes, tanto entre os com treinamento e sem treinamento, quanto entre patologistas e cientistas da computação (ver Figura 19). Analisando os intervalos de confiança do entre os grupos com e sem treinamento, pode-se notar visualmente

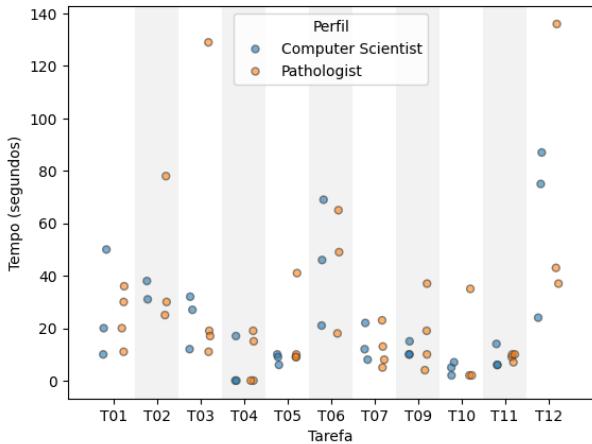
Figura 19 – Tempo de execução das tarefas bem-sucedidas



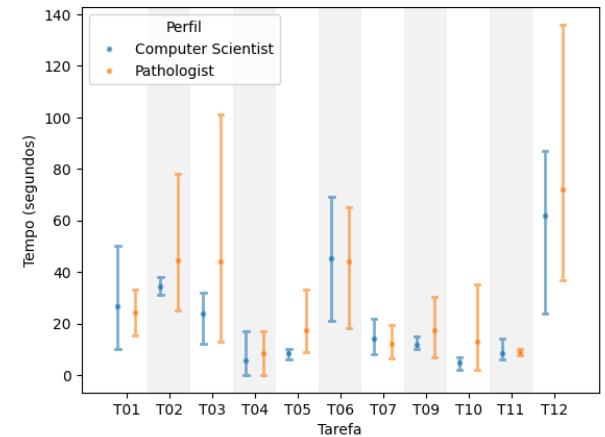
(a) Tempo por tarefa por grupo de treinamento



(b) Intervalo de confiança de (95%) tempo por tarefa dos participantes por grupo de treinamento



(c) Tempo por tarefa por perfil dos participantes



(d) Intervalo de confiança de (95%) tempo por tarefa por perfil dos participantes

Fonte: Elaborado pelo autora

que os participantes sem treinamento necessitam mais de tempo para executar as tarefas, mas um teste Mann-Whitney U comprovou que essa diferença não é significativa estatisticamente ( $z = 2.0$ ,  $p = 0.23$ ). Entre os perfis de participantes, é ainda mais evidente visualmente que não houve diferença, também comprovada pelo teste Mann-Whitney U ( $z = 5.0$ ,  $p = 0.86$ ).

A maioria das tarefas foi realizada sem dificuldades e sem a ocorrência de erro catastrófico. As tarefas consideradas mais fáceis pelos participantes foram tarefas de análise de rede. No caso da tarefa T08 - Descubra as características da doença *IGA NEPHROPATHY*, que nenhum dos participantes conseguiu realizar, há a hipótese de que a instrução da tarefa não estava clara o suficiente e confundiu os participantes, mesmo os que estavam no grupo com treinamento. Apesar de ser considerada uma tarefa difícil, havia a expectativa pelos avaliadores de que os participantes que estavam no grupo com treinamento conseguiram realizá-la. Com

exceção da Tarefa T08, todos os demais erros se concentraram no grupo sem treinamento, e a maior parte dos erros (4/5) foi cometida por patologistas.

O planejamento da avaliação foi avaliado por dois testes-piloto (como descrito na Seção 6.1.3.4). Todas as tarefas foram realizadas com sucesso pelo segundo teste-piloto, no entanto, não foi realizado nenhum teste-piloto com perfil de patologista, apenas com cientistas da computação. O teste-piloto com um participante patologista poderia ter resultado em ajustes no planejamento. Outra evidência de que o teste-piloto com patologista fez falta foi a ocorrência de mais erros na execução das tarefas pelos patologistas. Outra dificuldade na avaliação foi, apesar de terem sido recrutados 8 participantes, apenas 7 realizaram-na.

Como solução para melhorar o entendimento da ferramenta pelos usuários, pode-se investir na criação de *onboarding*<sup>26</sup> e uma *landing page*<sup>27</sup>. Por meio do *onboarding* o usuário, ao começar a usar o PS-iMAP, seria guiado por uma série de etapas para aprender como navegar no sistema, com o uso de *tours* interativos, tutoriais passo a passo ou dicas visuais. Por uma *landing page* sobre a ferramenta, poderiam ser apresentados o objetivo, as funcionalidades e a explicação de como funciona o PS-iMAP.

Concluindo-se a análise desta avaliação, todos os resultados dessa avaliação indicaram que os usuários entendem as informações apresentadas e que consideram o apoio computacional adequado na realização de suas tarefas. Os principais resultados obtidos foram: as informações apresentadas no PS-iMAP são relevantes para os usuários e eles não tiveram problemas críticos em operar a ferramenta. Entretanto, houve dificuldades no entendimento do objetivo da ferramenta. Em geral, essa avaliação mostrou que a visualização de rede fornece as informações de forma comprehensível para os participantes, tanto para patologistas quanto cientistas da computação, e foi o componente da ferramenta que mais chamou atenção.

### 6.3 Considerações Finais

Neste capítulo, descreveu-se a aplicação da abordagem proposta para realizar uma avaliação formativa, incluindo uma análise dos resultados produzidos. Com a entrega dos resultados, os *stakeholders* do PS-iMAP ficaram muito satisfeitos com os resultados do teste de

<sup>26</sup> O *onboarding* serve para introduzir e guiar novos usuários, clientes ou colaboradores em um ambiente, plataforma, serviço ou empresa. Seu principal objetivo é facilitar a adaptação inicial e assegurar que as pessoas entendam rapidamente como usar um sistema ou desempenhar suas funções.

<sup>27</sup> Uma *landing page* é uma página web projetada com o objetivo específico de converter visitantes. Ela é usada como parte de campanhas de marketing digital para captar a atenção dos usuários e incentivar uma ação imediata, como preencher um formulário, se inscrever em uma *newsletter*, baixar um material ou realizar uma compra.

usabilidade e iniciaram o desenvolvimento dos pontos de melhorias para a ferramenta.

Do ponto de vista desta pesquisa, a aplicação da abordagem proposta produziu lições aprendidas que serão abordadas no capítulo a seguir.

## 7 LIÇÕES APRENDIDAS

Todo o processo de avaliação do PS-iMAP descrito neste texto ilustrou alguns dos benefícios do método de pesquisa qualitativa, pois ele permitiu que fossem observados os comportamentos e necessidades dos usuários, resultando em um processo de avaliação mais eficaz e centrado no usuário. Foram observados grandes benefícios na aplicação da abordagem proposta. A combinação do teste de usabilidade com tarefas de análise de rede em condições que refletem a realidade proveu *insights* não apenas sobre a visualização de dados, mas também sobre a ferramenta como um todo.

Neste capítulo, são discutidas as lições aprendidas durante a execução do trabalho e reflexões acerca do método de avaliação aplicado e dos resultados reportado no Capítulo 6.

### 7.1 Métodos de avaliação de visualização de dados

A partir do levantamento da literatura, apesar de existirem recomendações para realizar as avaliações em condições que refletem a realidade, percebeu-se que a maioria das avaliações de redes multivariadas (GHONIEM *et al.*, 2005; KELLER *et al.*, 2006; JIANU *et al.*, 2014; OKOE *et al.*, 2018; REN *et al.*, 2019; NOBRE *et al.*, 2020; ABDELAAL *et al.*, 2022) não segue essas indicações.

Normalmente, os trabalhos são de caráter apenas quantitativo com usuários recrutados por ferramenta de *crowdsourcing*. A utilização desse tipo de ferramenta pode gerar alguns desafios como: recrutamento não direcionado, pois não há garantia de que o público-alvo esteja sendo representado da maneira adequada; falta de interação com os participantes para tirar eventuais dúvidas; e os participantes podem não estar tão engajados.

Além do mais, essas avaliações têm focado apenas nas comparações de abordagens das redes (diagrama nó-link e matriz de adjacência), não se importando com os recursos interativos que manipulam a visualização de dados como: zoom, filtros e acesso a detalhes dos dados, ou com a ferramenta em que a visualização está inclusa.

Em geral, há o foco exagerado na eficácia e eficiência na execução de tarefas, ou seja, há o ponto de vista de que as visualizações servem apenas para dar suporte às tarefas. Não importa o quanto habilmente a informação tenha sido codificada visualmente, idealmente todas as ferramentas de visualização de dados deveriam considerar outros fatores de usabilidade e experiência de uso em condições que refletem a realidade.

## 7.2 Condições que refletem a realidade

A avaliação com usuários reais, utilização de visualização real e avaliação da interação forneceram resultados não apenas sobre os fatores de usabilidade: eficácia e eficiência, como visto nos trabalhos relacionados sobre avaliação de redes multivariadas. Para criar uma ferramenta que realmente atenda às expectativas dos usuários, foi aplicada uma abordagem mais holística de avaliação considerando também a facilidade de aprendizado, a facilidade de uso e a percepção da ferramenta pelos usuários.

Ao realizar a avaliação com usuários reais foi possível verificar o interesse dos potenciais usuários e entender suas necessidades e dificuldades possibilitando ajustes na interface e na visualização de dados. Além de validar se as suposições e decisões de design realmente funcionam no contexto prático.

Ao usar uma visualização real com dados reais, a abordagem permitiu validar se a organização dos dados fazem sentido para os usuários e se, de fato, a visualização ajuda na tomada de decisões e compreensão da informação.

Por meio da análise de como os usuários interagem com a ferramenta, foi possível:

- Identificar como os usuários realmente utilizam a ferramenta e a visualização de dados.
- Detectar problemas que causam erros ou frustrações.
- Coletar dados que serão utilizados para guiar o design e desenvolvimento da ferramenta.
- Coletar *feedback* qualitativo dos usuários sobre sua experiência, dificuldades encontradas, e sugestões de melhorias.

A qualidade de uso é um aspecto fundamental em qualquer ferramenta com que os usuários interagem, ela depende da visão do usuário utilizando a ferramenta em condições que refletem a realidade. Portanto, faz-se necessário que as avaliações de redes multivariadas incluam essas condições e tarefas que direcionem a análise da visualização dentro de um determinado domínio.

## 7.3 Tarefas de análise de visualização de dados

A utilização de tarefas de análise de rede baseada em domínio e taxonomia ajudou a desenvolver evidências concretas a respeito da eficácia e eficiência da visualização de dados e a

avaliação da adequação da visualização para resolver o problema de domínio.

O designer de visualização deve dar atenção tanto para a tarefa de análise que será realizada pelos usuários, quanto para o tipo de dados, pois uma visualização com um conjunto de dados específico que serve bem para um tipo de tarefa pode não ser adequado para outras. No planejamento da avaliação houve a tentativa de formular as tarefas de análise da rede baseado apenas nas taxonomias (ignorando o domínio), e percebeu-se que as tarefas não faziam parte do contexto dos usuários, visto que o tipo de análise não produziria dados úteis. Por exemplo, uma tarefa de identificação que tenha como alvo a estrutura topológica de caminho (qual a doença em um caminho mais curto entre a nefropatia por IGA e a esclerose segmentar?) não faria sentido para a rede do PS-iMAP e seus usuários.

#### **7.4 Teste de usabilidade**

A aplicação do teste de usabilidade com tarefas de análise de rede em condições que refletem a realidade podem produzir ferramentas melhores e mais utilizáveis, validar a abordagem de visualização utilizada e, possivelmente, podem fornecer evidências da superioridade de uma abordagem de visualização para um contexto específico por meio da comparação.

Ao separar as tarefas em análise de rede e de interação com a ferramenta foi possível verificar separadamente os resultados das tarefas de visualização de dados e dos recursos interativos da ferramenta. Visto que, segundo Carpendale (2008), ao avaliar um software de visualização complexo, pode não estar claro se os resultados são devidos a uma técnica subjacente específica ou à solução geral do sistema.

O teste foi aplicado de maneira remota dependendo de uma conexão estável com a internet e ferramentas de videoconferência. A modalidade remota e moderada exigiu um tempo significativo para agendamento com os participantes e condução dos testes.

A etapa mais importante em qualquer pesquisa com usuários é o planejamento. O sucesso de uma pesquisa depende muito de um planejamento bem-feito, pois ele garante que a pesquisa produza resultados úteis, éticos e alinhado com os objetivos. Além disso, normalmente é a etapa mais demorada.

Apesar de possuir perfis de usuários bem distintos, a execução dos testes-piloto da avaliação do PS-iMAP foi executada apenas com um perfil de usuário. Para garantir que o teste final seja representativo e evitar problemas metodológicos durante a condução da avaliação, seria interessante realizar o teste-piloto com perfis diferentes de usuários.

Além do mais, durante o recrutamento dos participantes, se a avaliação for realizada com perfil muito específico de usuários, pode ser muito difícil obter seu tempo e uma amostra adequada.

A adaptação é essencial para enfrentar desafios ou imprevistos. Embora o planejamento seja uma parte crucial, nem todos os aspectos, principalmente envolvendo pessoas, podem ser antecipados. Um exemplo foi que no primeiro teste com patologista, o participante inicialmente respondeu as tarefas baseado em seus próprios conhecimentos, sendo necessário ajustar a maneira como a tarefa era solicitada.

Para coletar dados sobre a satisfação do usuário do PS-iMAP, foi realizada a entrevista pós-teste para coleta de *feedback* qualitativo e analisado todos as gravações para observar as reações e comentários dos usuários durante o teste. Porém seria mais eficiente se também tivesse sido realizado a aplicação de um questionário específico para quantificar as respostas.

Embora o teste de usabilidade possa ser aplicado em qualquer momento e com qualquer nível de fidelidade do protótipo, se o objetivo for avaliar a usabilidade de uma ferramenta de visualização no contexto de tarefas de análise de redes, é fundamental adotar uma abordagem mais robusta. Isso inclui a utilização de dados reais, a participação de usuários representativos do público-alvo, uma motivação realista e um protótipo que permita interações. Sem essas condições, a avaliação da usabilidade será limitada e poderá não gerar *insights* relevantes.

## 7.5 Considerações finais

Este capítulo apresentou as lições extraídas com base na pesquisa e nos resultados da aplicação da abordagem de avaliação proposta. Métodos de avaliação usados em visualização de dados, a criação de condições que refletem a realidade de uso do sistema avaliado, o cuidado na escolha das tarefas de rede e no planejamento do teste de usabilidade são fatores que influenciam diretamente na qualidade da avaliação.

## 8 CONCLUSÃO E TRABALHOS FUTUROS

Com a finalidade de tornar a experiência mais satisfatória e produtiva para os usuários de ferramentas de visualização de redes multivariadas, este trabalho investigou como utilizar o teste de usabilidade para avaliar esse tipo de ferramenta, visando tornar o teste de usabilidade relevante para pesquisadores de visualização de dados.

A abordagem metodológica proposta foi concebida após o levantamento de metodologias de avaliação de visualização de dados e de IHC, de redes multivariadas e de taxonomias de tarefas de análise. A metodologia foi aplicada em uma ferramenta web, fornecendo informações valiosas sobre a interação dos usuários com a ferramenta e com a visualização de dados, apresentando resultados promissores para a coleta de métricas quantitativas e dados qualitativos. O trabalho buscou inspirar e incentivar o uso de métodos de IHC para melhorias na usabilidade e foco no usuário na área de visualização de dados.

O trabalho procurou oferecer meios pelos quais seja possível responder à seguinte pergunta: “Como adaptar o teste de usabilidade para avaliar ferramentas com visualização de redes multivariadas de forma mais abrangente?” Como resultado, observou-se que a inclusão de tarefas de análise de redes em testes de usabilidade, realizadas em condições que refletem a realidade, proporciona *insights* valiosos não apenas sobre a visualização de dados, mas também sobre a ferramenta como um todo.

Ao longo do caminho, a pesquisa passou por limitações. Evidenciou-se a complexidade de avaliar ferramentas de visualização de dados com diferentes perfis de usuários, demandando um planejamento detalhado e tempo significativo para condução e análise dos resultados. Os objetivos foram atingidos dentro de uma série de limitações na aplicação da abordagem proposta. Elas são explicitadas a seguir:

- Aplicação da abordagem de avaliação em apenas uma ferramenta.
- Falta de metodologias de avaliação com etapas detalhadas, exemplificando a condução e os materiais utilizados;
- Para aplicação da avaliação, houve a necessidade da validação e aprovação constante dos *stakeholders* da ferramenta, visto que o contexto da ferramenta avaliada era específico;
- Demora na aprovação do projeto de pesquisa pelo comitê de ética, atrasando o início das avaliações;
- A necessidade de participantes com um perfil específico resultou em uma dispo-

nibilidade limitada. Ainda assim, a avaliação foi realizada com sete participantes relevantes que representaram adequadamente o público-alvo desejado.

- Falta de aplicação da abordagem proposta por terceiros;

A aplicação da abordagem de avaliação forneceu resultados detalhados sobre a usabilidade da ferramentas PS-iMAP, destacando aspectos como facilidade de uso, entendimento visual e tempo necessário para completar as tarefas analíticas. Também revelou onde os usuários encontram dificuldades ao usar a ferramenta, ajudando a direcionar melhorias e inovações futuras.

As principais contribuições da abordagem executada são listadas a seguir:

- Uma abordagem de avaliação com etapas detalhadas que possa ser utilizada por designers e desenvolvedores de ferramentas de visualização de redes multivariadas;
- Disponibilização dos materiais de apoio criados no planejamento da avaliação, contendo tudo o que foi utilizado neste trabalho;
- Uma discussão sobre avaliar outros fatores de usabilidade além da eficácia e da eficiência, para a melhoria de uma ferramenta de visualização de dados.

Futuramente, pode-se testar a abordagem para avaliações não moderadas permitindo uma coleta de dados em maior escala, pois a avaliação pode ser realizada simultaneamente por muitos usuários, sem a necessidade de um facilitador. Além disso, seria interessante aplicar a metodologia com outras redes multivariadas e outras áreas de domínio.

Neste trabalho, foi utilizada uma ferramenta paga para gravar a chamada de vídeo durante a avaliação. Trabalhos posteriores podem fazer um levantamento de ferramentas que podem ser utilizadas nesse propósito e também para as avaliações não moderadas. Durante o processo de desenvolvimento do trabalho, foi notado que a abordagem de incluir tarefas de análise de visualização de dados em métodos de IHC poderia ser utilizada para outros tipos de visualização. Dito isso, seria interessante aplicar a abordagem com outras visualizações.

Por fim, pretende-se fornecer um guia online para a condução da avaliação e disponibilizá-lo para que a metodologia de avaliação possa ser utilizada por outras pessoas para validação e melhorias na metodologia.

## REFERÊNCIAS

- ABDELAAL, M.; SCHIELE, N. D.; ANGERBAUER, K.; KURZHALS, K.; SEDLMAIR, M.; WEISKOPF, D. Comparative evaluation of bipartite, node-link, and matrix-based network representations. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, v. 29, n. 1, p. 896–906, 2022.
- ALPER, B.; RICHE, N.; RAMOS, G.; CZERWINSKI, M. Design study of linesets, a novel set visualization technique. **IEEE transactions on visualization and computer graphics**, IEEE, v. 17, n. 12, p. 2259–2267, 2011.
- AMAR, R. A.; STASKO, J. T. Knowledge precepts for design and evaluation of information visualizations. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, v. 11, n. 4, p. 432–442, 2005.
- American Journal of Kidney Diseases. **American Journal of Kidney Diseases**. 2024. Disponível em: <https://www.ajkd.org/>. Acesso em: 28 jul. 2024.
- APARICIO, M.; COSTA, C. J. Data visualization. **Communication design quarterly review**, ACM New York, NY, USA, v. 3, n. 1, p. 7–11, 2015.
- BARBOSA, S. D. J.; SILVA, B. d.; SILVEIRA, M. S.; GASPARINI, I.; DARIN, T.; BARBOSA, G. D. J. **Interação humano-computador e experiência do usuário**. [S. l.]: Auto publicação, 2021.
- BASTIEN, J. C. Usability testing: a review of some methodological and technical aspects of the method. **International journal of medical informatics**, Elsevier, v. 79, n. 4, p. e18–e23, 2010.
- BEVAN, N. Quality in use: Meeting user needs for quality. **Journal of systems and software**, Elsevier, v. 49, n. 1, p. 89–96, 1999.
- BIKAKIS, N. Big data visualization tools. **arXiv preprint arXiv:1801.08336**, 2018.
- BOLT, N.; TULATHIMUTTE, T. **Remote research: Real users, real time, real research**. [S. l.]: Rosenfeld Media, 2010.
- BORSCI, S.; MACREDIE, R. D.; BARNETT, J.; MARTIN, J.; KULJIS, J.; YOUNG, T. Reviewing and extending the five-user assumption: A grounded procedure for interaction evaluation. **ACM Trans. Comput.-Hum. Interact.**, Association for Computing Machinery, New York, NY, USA, v. 20, n. 5, nov 2013. ISSN 1073-0516. Disponível em: <https://doi.org/10.1145/2506210>.
- BRUNO, F. Rastros digitais sob a perspectiva da teoria ator-rede. **Revista Famecos**, v. 19, n. 3, p. 681–704, 2012.
- BUDIU, R. Quantitative vs. qualitative usability testing. **Nielsen Norman Group**, v. 1, 2017.
- CARD, S. K.; MACKINLAY, J. The structure of the information visualization design space. In: IEEE. **Proceedings of VIZ'97: Visualization Conference, Information Visualization Symposium and Parallel Rendering Symposium**. [S. l.], 1997. p. 92–99.
- CARPENDALE, S. Evaluating information visualizations. In: **Information visualization: Human-centered issues and perspectives**. [S. l.]: Springer, 2008. p. 19–45.

CHEN, C.-h.; HÄRDLE, W. K.; UNWIN, A. **Handbook of data visualization.** [S. l.]: Springer Science & Business Media, 2007.

COLLINS, C.; PENN, G.; CARPENDALE, S. Bubble sets: Revealing set relations with isocontours over existing visualizations. **IEEE transactions on visualization and computer graphics**, IEEE, v. 15, n. 6, p. 1009–1016, 2009.

Connected Papers. **Connected Papers | Find and explore academic papers.** 2024. Disponível em: <https://www.connectedpapers.com/>. Acesso em: 25 jul. 2024.

ELLIS, G.; DIX, A. An explorative analysis of user evaluation studies in information visualisation. In: **Proceedings of the 2006 AVI Workshop on BEyond Time and Errors: Novel Evaluation Methods for Information Visualization.** New York, NY, USA: Association for Computing Machinery, 2006. (BELIV '06), p. 1–7. ISBN 1595935622. Disponível em: <https://doi.org/10.1145/1168149.1168152>.

FILIPOV, V.; ARLEO, A.; MIKSCH, S. Are We There Yet? A Roadmap of Network Visualization from Surveys to Task Taxonomies. **Computer Graphics Forum**, v. 42, 2023. Disponível em: <https://onlinelibrary.wiley.com/doi/abs/10.1111/cgf.14794>.

GANSNER, E. R.; HU, Y.; KOBOUROV, S. Gmap: Visualizing graphs and clusters as maps. In: **IEEE. 2010 IEEE Pacific Visualization Symposium (PacificVis).** [S. l.], 2010. p. 201–208.

GHONIEM, M.; FEKETE, J.-D.; CASTAGLIOLA, P. On the readability of graphs using node-link and matrix-based representations: a controlled experiment and statistical analysis. **Information visualization**, SAGE Publications Sage UK: London, England, v. 4, n. 2, p. 114–135, 2005.

GOODMAN, E.; KUNIAVSKY, M.; MOED, A. **Observing the User Experience, Second Edition: A Practitioner's Guide to User Research.** 2nd. ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2012. ISBN 0123848695.

HADDADI, H. Human-data interaction. **Encyclopedia of Human Computer Interaction**, 2016.

HADDADI, H.; HOWARD, H.; CHAUDHRY, A.; CROWCROFT, J.; MADHAVAPEDDY, A.; MORTIER, R. Personal data: Thinking inside the box. **arXiv preprint arXiv:1501.04737**, 2015.

HARTSON, H. R.; CASTILLO, J. C.; KELSO, J.; NEALE, W. C. Remote evaluation: the network as an extension of the usability laboratory. In: **Proceedings of the SIGCHI conference on human factors in computing systems.** [S. l.: s. n.], 1996. p. 228–235.

HOLTEN, D.; WIJK, J. J. V. Force-directed edge bundling for graph visualization. In: **WILEY ONLINE LIBRARY. Computer graphics forum.** [S. l.], 2009. v. 28, n. 3, p. 983–990.

HUANG, W.; EADES, P.; HONG, S.-H. Measuring effectiveness of graph visualizations: A cognitive load perspective. **Information Visualization**, SAGE Publications Sage UK: London, England, v. 8, n. 3, p. 139–152, 2009.

Interaction Design Foundation - IxDF. **What is Usability?** 2016. Disponível em: <https://www.interaction-design.org/literature/topics/usability>. Acesso em: 20 ago. 2024.

- ISENBERG, P.; ZUK, T.; COLLINS, C.; CARPENDALE, S. Grounded evaluation of information visualizations. In: **Proceedings of the 2008 Workshop on BEyond time and errors: novel evaLuation methods for Information Visualization.** [S. l.: s. n.], 2008. p. 1–8.
- ISO 9241. ISO. 2010.
- ISO/IEC 9126-1. ISO. 2001.
- JIANU, R.; RUSU, A.; HU, Y.; TAGGART, D. How to display group information on node-link diagrams: An evaluation. **IEEE Transactions on Visualization and Computer Graphics**, IEEE, v. 20, n. 11, p. 1530–1541, 2014.
- JOYCE, A. Formative vs. summative evaluations. **Nielsen Norman Group World Leaders in Research-Based User Experience**, 2019.
- KARRAY, F.; ALEMZADEH, M.; SALEH, J. A.; ARAB, M. N. Human-computer interaction: Overview on state of the art. **International journal on smart sensing and intelligent systems**, De Gruyter Poland, v. 1, n. 1, p. 137, 2008.
- KELLER, R.; ECKERT, C. M.; CLARKSON, P. J. Matrices or node-link diagrams: which visual representation is better for visualising connectivity models? **Information Visualization**, SAGE Publications Sage UK: London, England, v. 5, n. 1, p. 62–76, 2006.
- KERREN, A.; PURCHASE, H. C.; WARD, M. O. Introduction to multivariate network visualization. In: SPRINGER. **Multivariate Network Visualization: Dagstuhl Seminar# 13201, Dagstuhl Castle, Germany, May 12-17, 2013, Revised Discussions.** [S. l.], 2014. p. 1–9.
- LAUBHEIMER, P. Beyond the nps: Measuring perceived usability with the sus, nasa-tlx, and the single ease question after tasks and usability tests. **Nielsen Norman Group**, 2018.
- LEE, B.; PLAISANT, C.; PARR, C. S.; FEKETE, J.-D.; HENRY, N. Task taxonomy for graph visualization. In: **Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization.** [S. l.: s. n.], 2006. p. 1–5.
- LUTNICK, B.; GINLEY, B.; GOVIND, D.; MCGARRY, S.; LAVIOLETTE, P.; YACOUB, R.; JAIN, S.; TOMASZEWSKI, J.; JEN, K.-Y.; SARDER, P. An integrated iterative annotation technique for easing neural network training in medical image analysis. **Nature Machine Intelligence**, v. 1, p. 112–119, 02 2019.
- MCCLOSKEY, M. Turn user goals into task scenarios for usability testing. **Nielsen Norman Group**, v. 12, 2014.
- MORAN, K. Usability testing 101. **Nielsen Norman Group**, v. 1, p. 12, 2019.
- MORTIER, R.; HADDADI, H.; HENDERSON, T.; MCAULEY, D.; CROWCROFT, J. Challenges & opportunities in human-data interaction. **University of Cambridge, Computer Laboratory**, 2013.
- MORTIER, R.; HADDADI, H.; HENDERSON, T.; MCAULEY, D.; CROWCROFT, J. Human-data interaction: The human face of the data-driven society. **arXiv preprint arXiv:1412.6159**, 2014.
- MUNZNER, T. **Visualization analysis and design.** [S. l.]: CRC press, 2014.

- NELSON, E. T.; STAVROU, A. Advantages and disadvantages of remote asynchronous usability testing using amazon mechanical turk. In: SAGE PUBLICATIONS SAGE CA: LOS ANGELES, CA. **Proceedings of the Human Factors and Ergonomics Society Annual Meeting**. [S. l.], 2011. v. 55, n. 1, p. 1080–1084.
- NIELSEN, J. **Usability engineering**. [S. l.]: Morgan Kaufmann, 1994.
- NIELSEN, J. **Why you only need to test with 5 users**. [S. l.]: Useit. com Alertbox, 2000.
- NIELSEN, J. Usability metrics. **Nielsen Norman Group**, v. 21, 2001.
- NIELSEN, J. Aspects of design quality. **Nielsen Norman Group**, 2008.
- NIELSEN, J. Thinking aloud: The #1 usability tool. **Nielsen Norman Group**, v. 16, 2012.
- NIELSEN, J. *et al.* Usability 101: Introduction to usability. Fremont, CA, 2012.
- NOBRE, C.; MEYER, M.; STREIT, M.; LEX, A. The state of the art in visualizing multivariate networks. v. 38, n. 3, p. 807–832, 2019.
- NOBRE, C.; WOOTTON, D.; CUTLER, Z.; HARRISON, L.; PFISTER, H.; LEX, A. revisit: Looking under the hood of interactive visualization studies. In: **Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems**. [S. l.: s. n.], 2021. p. 1–13.
- NOBRE, C.; WOOTTON, D.; HARRISON, L.; LEX, A. Evaluating multivariate network visualization techniques using a validated design and crowdsourcing approach. In: **Proceedings of the 2020 CHI conference on human factors in computing systems**. [S. l.: s. n.], 2020. p. 1–12.
- OKOE, M.; JIANU, R.; KOBOUROV, S. Node-link or adjacency matrices: Old question, new insights. **IEEE transactions on visualization and computer graphics**, IEEE, v. 25, n. 10, p. 2940–2952, 2018.
- OLIVEIRA, L.; CHAGAS, P.; DUARTE, A.; CALUMBY, R.; SANTOS, E.; ANGELO, M.; SANTOS, W. dos. Pathospotter: computational intelligence applied to nephropathology. In: **Innovations in Nephrology: Breakthrough Technologies in Kidney Disease Care**. [S. l.]: Springer, 2022. p. 253–272.
- PLAISANT, C. The challenge of information visualization evaluation. In: **Proceedings of the working conference on Advanced visual interfaces**. [S. l.: s. n.], 2004. p. 109–116.
- PREECE, J.; ROGERS, Y.; SHARP, H. **Design de interação**. [S. l.]: bookman, 2005.
- PRETORIUS, J.; PURCHASE, H. C.; STASKO, J. T. Tasks for multivariate network analysis. In: SPRINGER. **Multivariate Network Visualization: Dagstuhl Seminar# 13201, Dagstuhl Castle, Germany, May 12-17, 2013, Revised Discussions**. [S. l.], 2014. p. 77–95.
- REN, D.; MARUSICH, L. R.; O'DONOVAN, J.; BAKDASH, J. Z.; SCHAFFER, J. A.; CASSENTI, D. N.; KASE, S. E.; ROY, H. E.; LIN, W.-y. S.; HÖLLERER, T. Understanding node-link and matrix visualizations of networks: A large-scale online experiment. **Network Science**, Cambridge University Press, v. 7, n. 2, p. 242–264, 2019.
- ROCHA, H. V. D.; BARANAUSKAS, M. C. C. **Design e avaliação de interfaces humano-computador**. [S. l.]: Unicamp, 2003.

- ROSSI, R. A.; AHMED, N. K. The network data repository with interactive graph analytics and visualization. In: **AAAI**. [S. n.], 2015. Acessado em: 12/06/2022. Disponível em: <https://networkrepository.com>.
- RUBIN, J.; CHISNELL, D. **Handbook of usability testing: How to plan, design, and conduct effective tests**. [S. l.]: John Wiley & Sons, 2008.
- SADIKU, M.; SHADARE, A. E.; MUSA, S. M.; AKUJUOBI, C. M.; PERRY, R. Data visualization. **International Journal of Engineering Research And Advanced Technology (IJERAT)**, v. 2, n. 12, p. 11–16, 2016.
- SAKET, B.; ENDERT, A.; STASKO, J. Beyond usability and performance: A review of user experience-focused evaluations in visualization. In: **Proceedings of the Sixth Workshop on Beyond Time and Errors on Novel Evaluation Methods for Visualization**. New York, NY, USA: Association for Computing Machinery, 2016. (BELIV '16), p. 133–142. ISBN 9781450348188. Disponível em: <https://doi.org/10.1145/2993901.2993903>.
- SANTAELLA, L.; KAUFMAN, D. Os dados estão nos engolindo? **Civitas-Revista de Ciências Sociais**, SciELO Brasil, v. 21, p. 214–223, 2021.
- SANTOS, J.; SILVA, R.; OLIVEIRA, L.; SANTOS, W.; ALDEMAN, N.; DUARTE, A.; VERAS, R. Glomerulosclerosis detection with pre-trained cnns ensemble. **Computational Statistics**, Springer, v. 39, n. 2, p. 561–581, 2024. Disponível em: <https://doi.org/10.1007/s00180-022-01307-3>.
- SCHADE, A. Remote usability tests: moderated and unmoderated. **Evidence-Based User Experience Research, Training, and Consulting**. NN/g Nielsen Norman Group, 2013.
- SCHMIDT, A.; ALT, F.; MÄKELÄ, V. Evaluation in human-computer interaction—beyond lab studies. In: **Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems**. [S. l.: s. n.], 2021. p. 1–4.
- SEDLMAIR, M.; MEYER, M.; MUNZNER, T. Design study methodology: Reflections from the trenches and the stacks. **IEEE transactions on visualization and computer graphics**, IEEE, v. 18, n. 12, p. 2431–2440, 2012.
- SHARPE, H.; ROGERS, Y.; PREECE, J. **Interaction design: beyond human-computer interaction 2nd ed.** [S. l.]: John Wiley Sons, 2007.
- SHNEIDERMAN, B.; PLAISANT, C. Strategies for evaluating information visualization tools: multi-dimensional in-depth long-term case studies. In: **Proceedings of the 2006 AVI workshop on BEyond time and errors: novel evaluation methods for information visualization**. [S. l.: s. n.], 2006. p. 1–7.
- SINHA, G.; SHAHI, R.; SHANKAR, M. Human computer interaction. In: **IEEE. 2010 3rd International Conference on Emerging Trends in Engineering and Technology**. [S. l.], 2010. p. 1–4.
- THOMPSON, K. E.; ROZANSKI, E. P.; HAAKE, A. R. Here, there, anywhere: remote usability testing that works. In: **Proceedings of the 5th conference on Information technology education**. [S. l.: s. n.], 2004. p. 132–137.

- VALIATI, E. R.; PIMENTA, M. S.; FREITAS, C. M. A taxonomy of tasks for guiding the evaluation of multidimensional visualizations. In: **Proceedings of the 2006 AVI workshop on Beyond time and errors: novel evaluation methods for information visualization.** [S. l.: s. n.], 2006. p. 1–6.
- VICTORELLI, E. Z.; REIS, J. C. D.; HORNUNG, H.; PRADO, A. B. Understanding human-data interaction: Literature review and recommendations for design. **International journal of human-computer studies**, Elsevier, v. 134, p. 13–32, 2020.
- WARE, C. **Visual thinking for design.** [S. l.]: Elsevier, 2010.
- WEHREND, S.; LEWIS, C. A problem-oriented classification of visualization techniques. In: **IEEE. Proceedings of the First IEEE Conference on Visualization: Visualization90.** [S. l.], 1990. p. 139–143.
- WHITENTON, K. Unmoderated user tests: how and why to do them. **Nielsen Norman Group.** URL: <https://www.nngroup.com/articles/unmoderated-usability-testing/> [accessed 2021-11-09], 2019.
- YI, J. S. Qndreview: Read 100 chi papers in 7 hours. In: **CHI'14 Extended Abstracts on Human Factors in Computing Systems.** [S. l.: s. n.], 2014. p. 805–814.
- YIN, R. K. **Case study research.** [S. l.]: Thousand Oaks, CA: Sage, 2014.
- YOGHOURDJIAN, V.; YANG, Y.; DWYER, T.; LAWRENCE, L.; WYBROW, M.; MARRIOTT, K. Scalability of network visualisation from a cognitive load perspective. **IEEE Transactions on Visualization & Computer Graphics**, IEEE Computer Society, Los Alamitos, CA, USA, v. 27, n. 02, p. 1677–1687, feb 2021. ISSN 1941-0506.

**APÊNDICE A – TERMOS DE CONSENTIMENTO LIVRE E ESCLARECIDO**

# TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Prezado(a) participante,

Você está sendo convidado(a) pela pesquisadora Bruna Neves de Oliveira, a participar da pesquisa intitulada “PS-T: Plataforma de testes da Rede de Pesquisas PathoSpotter”, sob a orientação da Profa. Dra. Emanuele Santos e coordenação do Prof. Dr. Washington Santos. O objetivo geral desta pesquisa é avaliar o uso e identificar problemas de interação com as ferramentas da plataforma Pathospotter. Essas ferramentas visam auxiliar patologistas em suas atividades diagnósticas e de pesquisa.

Sua participação no referido estudo é voluntária e será no sentido de identificar problemas e dificuldades durante o uso da ferramenta PS-iMAP. O foco da pesquisa é a qualidade e facilidade de uso da ferramenta e não visa conclusões sobre os usuários. O tempo estimado para a participação da pesquisa é em torno de 30 a 40 minutos. Pedimos seu consentimento para a realização e gravação de uma entrevista pré e pós a avaliação, e também a gravação da tela do computador durante o uso da ferramenta. As etapas referentes à sua participação serão as seguintes:

1. Preencher um formulário sociodemográfico e de coleta de dados sobre a sua experiência prévia com conceitos de patologia e ciência da computação.
2. Receber orientações sobre as principais funcionalidades da ferramenta proposta.
3. Usar a ferramenta para responder algumas perguntas de acordo com o roteiro fornecido pelo pesquisador.
4. Preencher um formulário de avaliação sobre a experiência de uso da ferramenta.
5. Participar de uma entrevista informal sobre a sua opinião como usuário(a) e sobre o que você acha necessário melhorar na ferramenta proposta;
6. Algumas etapas do estudo são gravadas em vídeo.

Como se trata de uma pesquisa não invasiva e não designada a tirar conclusões sobre pessoas, os riscos envolvidos são mínimos, listaremos alguns aqui: pode haver desconfortos decorrentes da participação na pesquisa como cansaço ao responder os questionários, não se sentir à vontade para expor a opiniões durante a realização das entrevistas ou durante gravações de áudio e vídeo, ou ao não conseguir realizar alguma das atividades. Caso isso ocorra durante sua participação, ou se você se sentir desconfortável ou prejudicado de alguma forma, interrompa a sua participação e entre em contato com os pesquisadores ou com o coordenador do trabalho nos contatos fornecidos abaixo. Você terá

direito a indenização no caso de danos decorrentes do estudo, bem como ao ressarcimento de despesas e compensação material decorrentes do estudo.

Você pode se recusar a participar do estudo, ou retirar seu consentimento a qualquer momento, sem precisar justificar e sem qualquer prejuízo para você.

Os dados coletados durante as entrevistas e a avaliação destinam-se estritamente a atividades de análise da aplicação. Os resultados serão divulgados em artigos científicos, congressos e seminários. A divulgação desses resultados pauta-se no respeito à sua privacidade, e o seu anonimato será preservado em quaisquer documentos que elaborarmos.

Não há compensações diretas a você por participar da pesquisa. Contudo, ao participar, você estará aprendendo sobre desenvolvimentos na área de patologia computacional e contribuindo geração de produtos computacionais que ajudarão a melhorar a assistência médica e a democratização de diagnóstico aos quais muitas pessoas não têm acesso.

Os responsáveis estarão à sua disposição para qualquer esclarecimento que você considere necessário em qualquer etapa da pesquisa. Caso você tenha dúvidas ou perguntas adicionais entre em contato com: [brunaneves@alu.ufc.br](mailto:brunaneves@alu.ufc.br), [emanuele@dc.ufc.br](mailto:emanuele@dc.ufc.br) e [washington.conrado@ufba.br](mailto:washington.conrado@ufba.br).

O estudo foi aprovado pelo Comitê de Ética em Pesquisa do Instituto Gonçalo Moniz - FIOCRUZ, CAAE 77471724.2.0000.0040. Em caso de dúvida quanto à condução ética do estudo, entre em contato com: [cep.igm@fiocruz.br](mailto:cep.igm@fiocruz.br) ou (71) 3176 - 2285.

---

\* Indica uma pergunta obrigatória

1. E-mail \*

---

2. Nome \*

---

3. Diante do exposto, declaro que fui devidamente esclarecido(a) e dou o meu \*  
consentimento para participar da pesquisa e para a publicação dos resultados.

*Marque todas que se aplicam.*

Eu li e concordo com esses termos e condições.

---

Este conteúdo não foi criado nem aprovado pelo Google.

**Google Formulários**

## APÊNDICE B – ENTREVISTA ESTRUTURADA PRÉ-TESTE

### Entrevista estruturada pré-teste

(Pergunta 12 é apenas para os médicos patologistas)

1. Qual seu nome?
2. Qual a sua idade?
3. Qual o seu gênero?
4. Qual é o seu nível de escolaridade?
5. Qual a sua área de formação/atuação?
6. Quanto tempo de experiência você possui na sua área de formação? Se você ainda estuda, qual é o seu curso e período?
7. Com que frequência você usa o computador?
8. Com que frequência você utiliza a internet?
9. Qual navegador você utiliza (Por exemplo: Google Chrome, Firefox, Opera, Explorer, Safari)?
10. Você sabe o que é uma estrutura histológica?
11. Você já visitou algum site para visualizar alguma estrutura histológica? Qual(is)?
12. Qual a sua dificuldade na caracterização desse tipo de estrutura? Em uma escala de 1 a 3, em que 1 é mínimo e 3 é o máximo.

## APÊNDICE C – CENÁRIO COM TAREFAS PARA PATOLOGISTAS

### **Cenário geral**

Marcos tem 37 anos, é médico patologista e nas horas livres gosta de andar de bicicleta e de fazer ilustrações médicas relacionadas a nefropatologia e postá-las no Instagram. Recentemente, Marcos começou a colaborar com Ana, uma pesquisadora em Ciência da Computação, para aplicar técnicas de Machine Learning em imagens de lesões em biópsias renais para auxiliar no diagnóstico de doenças renais. Ana, recentemente o recomendou uma ferramenta online chamada PS-iMAP para a visualização das lesões e das técnicas de Patologia Computacional mais utilizadas.

---

### **Parte 1 - Exploração livre**

Por favor, realize as tarefas a seguir como se fosse Marcos:

1. Realize o login com os seguintes dados (apresentar os dados)
  2. Explore a ferramenta livremente e relate suas descobertas (Há algo surpreendente ou particularmente interessante?)
- 

### **Parte 2 - Interação com a visualização de dados**

Ana, sempre interessada em debater sobre as classificações de lesões, pede para Marcos utilizar o PS-iMAP para que possam chegar a um consenso sobre que técnicas utilizar para classificar algumas lesões renais. Marcos, que gostou da sua experiência inicial com a ferramenta, tenta utilizá-la

---

Por favor, realize as tarefas a seguir como se fosse Marcos:

1. Localize a lesão hypercellularity (endocapillary)
2. Encontre todas doenças ligadas à lesão hypercellularity (endocapillary)
3. Selecione a lesão sclerosis, segmental
4. Explore a rede de maneira focada
5. Liste as doenças ligadas à lesão sclerosis, segmental
6. Encontre a lesão com o maior número de doenças
7. Localize e descubra como acessar as características da doença iga nephropathy

---

### Parte 3 - Interação com outros recursos da ferramenta PS-iMAP

Marcos achou interessante a ferramenta que Ana apresentou, mas depois de encontrar as doenças ligadas à lesão, ele interage com outros recursos disponíveis no PS-iMAP que o ajudarão na sua pesquisa.

---

Por favor, realize as tarefas a seguir como se fosse Marcos:

1. Localize a lesão sclerosis, segmental
2. Descubra quantas imagens estão associadas à lesão sclerosis, segmental
3. Identifique quantos resultados de trabalhos relacionados à lesão sclerosis, segmental foram encontrados e indique os que são de acesso aberto
4. Quais as atividades do patologista que já foram automatizadas nos primeiros 3 trabalhos relacionados à lesão sclerosis, segmental?

## **APÊNDICE D – CENÁRIO COM TAREFAS PARA CIENTISTA DA COMPUTAÇÃO**

### **Cenário geral**

Ana tem 33 anos, é recém-doutora em Ciência da Computação e nas horas livres gosta de ouvir podcasts e participar da comunidade do Stack Overflow respondendo perguntas não resolvidas de Machine Learning. Recentemente, Ana começou a colaborar com Marcos, um médico patologista interessado em aplicar técnicas de Machine Learning em imagens de lesões em biópsias renais para auxiliar no diagnóstico de doenças renais. Marcos, recentemente a recomendou uma ferramenta online chamada PS-iMAP para a visualização das lesões e das técnicas de Patologia Computacional mais utilizadas.

---

### **Parte 1 - Exploração livre**

Por favor, realize as tarefas a seguir como se fosse a Ana:

1. Realize o login com os seguintes dados (apresentar os dados)
  2. Explore a ferramenta livremente e relate suas descobertas (Há algo surpreendente ou particularmente interessante?)
- 

### **Parte 2 - Interação com a visualização de dados**

Marcos, sempre interessado em debater sobre as classificações de lesões, pede para Ana utilizar o PS-iMAP para que possam chegar a um consenso sobre que técnicas utilizar para classificar algumas lesões renais. Ana, que gostou da sua experiência inicial com a ferramenta, tenta utilizá-la.

---

Por favor, realize as tarefas a seguir como se fosse a Ana:

1. Localize a lesão hypercellularity (endocapillary)
2. Encontre todas doenças ligadas à lesão hypercellularity (endocapillary)
3. Selecione a lesão sclerosis, segmental
4. Explore a rede de maneira focada
5. Liste as doenças ligadas à lesão sclerosis, segmental
6. Encontre a lesão com o maior número de doenças
7. Localize e descubra como acessar as características da doença iga nephropathy

---

### Parte 3 - Interação com outros recursos da ferramenta PS-iMAP

Ana achou interessante a ferramenta que Marcos apresentou, mas depois de encontrar as doenças ligadas à lesão, ela interage com outros recursos disponíveis no PS-iMAP que a ajudarão na sua pesquisa.

---

Por favor, realize as tarefas a seguir como se fosse a Ana:

1. Localize a lesão sclerosis, segmental
2. Descubra quantas imagens estão associadas à lesão sclerosis, segmental
3. Identifique quantos resultados de trabalhos relacionados à lesão sclerosis, segmental foram encontrados e indique os que são de acesso aberto
4. Qual(is) a(s) principal(is) técnica(s) computacional(is) utilizada(s) nos primeiros 3 trabalhos relacionados à lesão sclerosis, segmental?

## APÊNDICE E – ENTREVITA SEMIESTRUTURADA PÓS-TESTE

### Entrevita semiestruturada pós-teste

1. Qual a sua opinião sobre a ferramenta PS-iMAP?
2. Qual foi a tarefa mais difícil?
3. Qual foi a tarefa mais fácil?
4. Encontrou algum termo que não conhecia?
5. Você teve alguma dúvida sobre a ferramenta?
6. Quais foram as suas dificuldades ou desconfortos ao realizar as tarefas?
7. Quais são as suas sugestões para a realização das tarefas de maneira mais fácil?
8. Você sugere alguma outra funcionalidade que poderia ser realizada com a ferramenta?
9. Qual o seu interesse em usar essa ferramenta no dia-a-dia?

## APÊNDICE F – ROTEIRO

**Objetivo:** Analisar apropriação da tecnologia para julgar se o uso do sistema oferece aos usuários o apoio computacional adequado e identificar problemas na interação e na interface.

**Perfil dos participantes:** Médicos patologistas e cientistas da computação

**Equipamentos:** Computador.

**Infraestrutura de software:** Navegador para ligação de vídeo, software para gravação de tela, conexão com a internet.

**Material de apoio:**

- Termo de consentimento
- Cenários com tarefas para os médicos patologistas
- Cenários com tarefas para os cientistas da computação

**Instruções para o ambiente de teste:**

Pedir para o participante compartilhar a tela durante a execução das tarefas e abrir no navegador o cenário e o PS-iMAP.

**Atividades iniciais:**

- Dar as boas-vindas e agradecer a participação
- Pedir para o participante abrir no link do cenário com as tarefas
- Iniciar a gravação da chamada
- Explicar o procedimento da avaliação e sobre o projeto (texto disponível nos materiais de apoio e abaixo)
- Ler o termo de consentimento e esperar a leitura (se o participante preferir)
- Perguntar se tem alguma dúvida sobre o termo
- Fornecer as instruções:
  - Caso precise, pode pedir auxílio.
  - Caso deseje, pode desistir da tarefa.
- Apresentar a introdução em vídeo sobre a ferramenta (para os participantes com treinamento) e perguntar se tem alguma dúvida sobre o vídeo

**Sobre o projeto:**

O PS-iMAP é uma ferramenta com associações de lesões e doenças visando melhorar a comunicação e a colaboração entre cientistas da computação e patologistas no desenvolvimento de projetos e parcerias ligados às duas áreas.

A Patologia Computacional surgiu como uma área de estudo que tem por finalidade integrar medicina e ciência da computação. No entanto, pesquisadores de ambas as áreas expericiam dificuldade na obtenção de informação do estado da arte da outra área e de comunicar seus próprios conhecimentos em uma equipe multidisciplinar. Neste contexto, ferramentas que auxiliem nessa troca de conhecimentos podem fomentar a comunicação e a colaboração entre esses pesquisadores.

O presente estudo tem como objetivo geral avaliar a usabilidade e o entendimento da ferramenta PS-iMAP.

**APÊNDICE G – ANÁLISE DE CONTEÚDO**

Participante	Tópico	Comentário
P1	Opinião	"Visualização como se fosse um estilo de grafo, achei bem legal que a gente consegue ver a dimensão da lesão, as conexões com os nós e as doenças também."
P1	Opinião	"O modo focal é bem legal. Como falou no vídeo, pode ser que ele fique sobreposto né? Não sei se poderia melhorar essa questão aqui de ver mais espaçado"
P1	Opinião	"Outra coisa que achei bem legal foi essa parte lateral de apresentar os artigos, tem os open access e os closed access que você pode visualizar."
P1	Opinião	"Isso [trabalhos relacionados] me lembra um pouco do github que você coloca os códigos e consegue ver outras informações"
P1	Sugestão	"Acho que talvez pudesse ter [a função] de código associado a esse artigo, a base de dados..."
P1	Sugestão	"Também poderia adicionar o fator de impacto dele, se ele é A1, A2"
P1	Sugestão	Sugeriu um filtro por corante
P1	Descoberta	Percebeu que ao buscar na rede, já filtra as imagens
P1	Sugestão	Sugeriu uma marcação na imagem já mostrando onde tá a lesão.
P1	Descoberta	Notou que precisava clicar na lupa para realizar a busca
P1	Sugestão	Sugeriu colocar na descrição da imagem o corante associado
P1	Dúvida	Perguntou se existia a função de buscar por imagem para ver a lesão associada
P1	Sugestão	Sugeriu padronizar a largura e a altura das imagens
P2	Descoberta	"Pode ter uma lesão sem nenhuma doença"
P2	Descoberta	"No caso, o tamanho do nó é [correspondente] ao tamanho de conexões"
P2	Opinião	"[tem algo que você achou interessante?] Acho que essa ideia de agrupar lesão e doença, é bem legal pra ver lesões que se repetem em várias doenças"
P2	Opinião	"É legal também ver esses casos mais isolados, né, ... que não tem nenhuma conexão."
P2	Opinião	"Acho que também poder ver uma imagem dá um overview bem legal sobre o que é"
P3	Opinião	"Essa visualização aqui não tá tão clara para mim"
P3	Descoberta	Percebeu que o nó da rede influência nas imagens mostradas a baixo.
P3	Dúvida	Perguntou como resetar a rede, mas na verdade queria saber como limpar os filtros da imagem
P3	Sugestão	Sugeriu que o campo de busca por lesão ou doença fosse no topo da tela
P3	Sugestão	Falou sobre carregar todas as imagens de uma só vez, citou a funcionalidade de borda infinita
P4	Observação	Ficou apontando trabalhos relacionados que achava que não fazia sentido.
P4	Observação	Perguntou se os artigos eram só para o pessoal de computação.
P4	Observação	Olhou várias imagens de lesões e foi falando se existia a lesão na imagem.
P4	Opinião	Falou que certas imagens estavam lindas
P4	Opinião	Explorou o modo focal e disse que ficou legal.
P4	Sugestão	Deu a sugestão: "criar uma aba aqui em baixo [da imagem da galeria de imagens aberta] fazer a descrição do que tem na lâmina, mas de resto ok"
P4	Opinião	"Ah eu gostei, é fácil de navegar"
P5	Sugestão	"Eu vi que ele realmente carrega todas as imagens né, se eu for descendo, não sei se isso pesa o site um pouco. Mas geralmente quando tem sites que mostram muitas coisas a gente vai apertando em "mostre mais"
P5	Opinião	"A primeira funcionalidade aqui foi explicada no vídeo né?"
P5	Opinião	"Essa rede ficou bem interessante mesmo"
P5	Descoberta	"Massa né? mostra uma lâmina aqui [card de detalhes] e o resultado dos artigos né?"
P5	Dúvida	Ficou com dúvida se os arquivos são atualizados automaticamente e de onde vêm.
P5	Opinião	Falou que seria melhor se os nós se distribuissem melhor já que ficam uns em cima de outros no modo focal, e por causa disso, disse que preferiria não utilizar o modo focal.
P5	Sugestão	"Seria massa também colocar o corte e a coloração que foi utilizada aqui porque eu vi que não foi padronizada, uns foram HE e outros foram ... sei lá diferente"
P5	Sugestão	"Não aparece nada aqui [card de detalhes] né quando é a doença selecionada? Seria legal se tivesse um resuminho igual aqui [lesão selecionada]"

P6	Opinião	"Euacheimuitomassa, o layout é muito legal... essa coisa da interconexão"
P6	Observação	Não deu scroll na página, então acabou não vendo a parte das imagens.
P7	Descoberta	"A sensação que eu tenho é como se eu tivesse mexendo uma metilação, como se fosse um processo de metilação, que é quando a gente agrupa algumas situações diante de algumas doenças, tipo eu vejo nefropatologia também né? então a sensação só vendo a imagem e não mexendo em muita coisa, é... eu acho que é como se a gente agrupasse as doenças e achasse os padrões daquelas doenças de acordo com o gráfico que a gente tá vendo aqui"
P7	Descoberta	"Ai você clica e aparece a imagem relacionada aquilo e o artigo relacionado aquela doença, então você consegue agrupar de acordo com a patologia..."
P7	Opinião	"Talvez seja uma ferramenta mais fácil para você conseguir dar um diagnóstico mais rápido, você bate o olho na figurinha e vem aqui no agrupamento da lesão e vai no artigo para descrever tudo o que tem relacionado aquela lesão, foi isso mais ou menos que eu entendi mas eu tô mexendo aqui ainda"
P7	Descoberta	"Ai, os em amarelo são tudo que tem de lesão e em azul é todas as doenças relacionadas aquela lesão"
P7	Opinião	"Muito bom"
P7	Opinião	"Ai tem imagens de todas as lesões, muito bom, muito legal"
P7	Opinião	"Acho que vai ajudar muito inclusive, porque hoje eu tive um caso na minha rotina como patologista que eu nunca tinha visto aquela lesão e quando eu procurei na internet eu tive muita dificuldade de achar uma imagem que tivesse... o patologista trabalha é como se fosse um quebra-cabeça, a gente vai juntando as peças e o raciocínio em cima daquilo e ai quando a gente raciocina que pode ser aquilo, a gente vai procurar figuras que pareçam com aquilo que a gente pensou. Eu acho que essa ferramenta vai ser muito prática porque a gente pensa por exemplo que é uma lesão mínima eu venho aqui embaixo já sei onde tá a lesão mínima e vejo todas as alterações que eu posso encontrar e eu bato a figurinha e já penso: não, é isso mesmo."
P7	Opinião	"Acho que vai ser bem legal a ferramenta. Muito bom".
P7	Descoberta	"E os artigos todos são de livre acesso?"
P7	Opinião	"E isso vai ficar disponível para gente? porque eu já fiquei interessada"
P7	Opinião	"Achei muito interessante, gostei muito"
P7	Opinião	"Acho que se fizesse isso de todas as áreas vai facilitar a rotina do patologista, tudo bem que assim, a quantidade de imagem é absurda e para você fazer isso, você precisa ter um banco muito grande de imagens, né? mas algumas pessoas que já fazem essas fotos, porque o dr Washington sempre fotografa os casos dele, então, facilita muito pra quem já fotografa"
P7	Dúvida	"Agora uma coisa que eu fiquei na dúvida é... eu já cliquei aqui na imagem né? eu sei que é uma tireoidização tubular como eu faço para voltar para saber onde tá essa tireoidização tubular, não tem como eu clicando aqui (na imagem do card descritivo) aparecer lá (na rede), não?"

**ANEXO A – PARECER DO COMITÊ DE ÉTICA EM PESQUISA**

**CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FOICRUZ/BA**



**PARECER CONSUBSTANCIADO DO CEP**

**DADOS DO PROJETO DE PESQUISA**

**Título da Pesquisa:** PS-T: Plataforma de testes da Rede de Pesquisas PathoSpotter

**Pesquisador:** WASHINGTON LUIS CONRADO DOS SANTOS

**Área Temática:**

**Versão:** 2

**CAAE:** 77471724.2.0000.0040

**Instituição Proponente:** Centro de Pesquisas Gonçalo Moniz - CPqGM/ FIOCRUZ/ BA

**Patrocinador Principal:** Financiamento Próprio

**DADOS DO PARECER**

**Número do Parecer:** 6.812.283

**Apresentação do Projeto:**

Informações contidas nos campos Apresentação do Projeto, Objetivo da Pesquisa e Avaliação dos Riscos e Benefícios foram retiradas do arquivo PB\_INFORMAÇÕES\_BÁSICAS\_DO\_PROJETO.

Parecer de resposta às pendências emitidas no parecer consubstanciado em 29 de Março de 2024.

**1. Introdução:**

A última década tem observado um importante avanço na área da patologia computacional. Sistemas têm sido desenvolvidos com capacidade de identificar e delinear estruturas em imagens histológicas, reconhecer e classificar estruturas normais e lesões. Apesar desses avanços poucas ferramentas têm chegado ao estágio de avaliação clínica. Um dos fatores limitantes para esse importante passo, no desenvolvimento das ferramentas de patologia computacional é a submissão dessas ferramentas a testes de uso, para identificação de potenciais falhas ou erros de comunicação com o usuário. Há quase 10 anos foi criada a rede PathoSpotter de patologia computacional. O objetivo dessa rede é produzir sistemas de inteligência artificial que auxiliem o patologista na sua rotina diagnóstica, estudantes e médicos residentes em sua formação em patologia e especialização na área, contribuam com a

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

# CENTRO DE PESQUISAS GONÇALO MONIZ - FIOCRUZ/BA



Continuação do Parecer: 6.812.283

pesquisa de novas estratégias terapêuticas e na integração de dados para composição de terapêutica mais precisas para grupos individualizados de pacientes. Ao longo desses anos de trabalho da rede PathoSpotter, algumas ferramentas de patologia computacional foram criadas pela rede e disponibilizada para uso por patologistas e outros indivíduos interessados da área biomédica. Progressivamente sentimos a necessidade de obter dados que nos ajudem a conhecer melhor a funcionalidade dessas ferramentas e identificar erros sistemáticos e ocasionais de desempenho. Dessa forma propomos neste projeto a criação de uma plataforma de testes sistemáticos das ferramentas do PathoSpotter que possibilite a coleta de informações que ajudem a aprimorar a performance dessas ferramentas. A execução deste projeto é fundamental para a criação de ferramentas mais robustas e seguras para os usuários e pacientes e que auxiliem a migração dessas ferramentas para as etapas clínicas de testagem.

## 2. Hipótese:

A submissão de sistemas a serem usados em patologia computacional ao teste de uso é etapa necessária para correção de erros de performance e comunicação com os usuários.

## 3. Metodologia Proposta:

a) Plataformas: Serão constituídas a partir de todos os sistemas computacionais de análise de imagens histológicas criadas no projeto PathoSpotter. Até o momento esses sistemas lidam com identificação, segmentação e classificação de estruturas normais e lesões em imagens histológicas do rim e educação em patologia computacional (por favor, acessem a lista completa de publicações referentes a esses sistemas em (<https://pathospotter.bahia.fiocruz.br/publications>)). Esses sistemas serão reunidos em diferentes plataformas de uso, com interfaces apropriadas para interação com usuários das áreas afins.

b) Participantes: Indivíduos adultos de nível de graduação ou graduados nas diferentes áreas do saber. Estima-se que cerca de 1000 indivíduos serão arrolados nesses testes no período de 5 anos do projeto.

c) Modelo do teste: Os indivíduos que aceitarem participar dos testes serão direcionados para as

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

# CENTRO DE PESQUISAS GONÇALO MONIZ - FIOCRUZ/BA



Continuação do Parecer: 6.812.283

plataformas presencial ou virtualmente, receberão informações detalhadas sobre a plataforma e o teste e, caso ainda queiram participar, assinarão um TCLE. Quando necessário, por exemplo no teste de plataformas de classificação de lesões, uma página de aviso sobre leitura e compromisso obrigatórios sobre os limites e responsabilidade no uso do resultado dos testes será exibida. O indivíduo deverá então encaminhado para as seguintes etapas do teste da plataforma:

- 1) Exploração livre da ferramenta computacional;
- 2) Responder algumas questões através do uso da plataforma;
- 3) Responder um formulário sobre as dificuldades e facilidades que encontrou ao usar a plataforma. Todas as etapas do teste serão registradas através da supervisão direta do pesquisador responsável pela aplicação do teste e pelo uso de ferramentas computacionais de registro de atividades.

d) Variáveis examinadas: O teste aplica-se à performance da plataforma, não das pessoas. Dessa forma não interessam a este projeto inferências sobre os indivíduos participantes, mas sobre o quanto as plataformas permitem conforto de uso e respostas ao fim para qual foram criadas. É necessário, então, obter informações gerais sobre a adequação da população de uso da plataforma, durante os testes. Serão colhidas por questionário as seguintes variáveis demográficas: sexo, idade, nível de formação acadêmica, endereço e outras variáveis que permitam o conhecimento dessa população de usuários. Adicionalmente serão colhidas variáveis indicadoras de usabilidade da ferramenta: facilidade de manuseio, interatividade, facilidade de encontrar intuitivamente funções, buscas na plataforma resultando em sucesso ou frustração, respostas a questões sobre relativas ao uso da ferramenta. Representação e análise dos resultados: Os dados coletados serão representados em tabelas e gráficos. Serão analisados como variáveis categóricas, ordinais ou numéricas, serão submetidos a análises de concordância, correlações e comparações entre grupos. Dado a natureza dos dados muito das análises utilizarão testes estatísticos não paramétricos, sendo, contudo, utilizados testes paramétricos sempre que possível.

## Objetivo da Pesquisa:

Objetivo Primário:

Adequar plataformas computacionais para uso em pesquisa e diagnóstico.

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FIOCRUZ/BA



Continuação do Parecer: 6.812.283

**Objetivo Secundário:**

- 1) Submeter plataformas geradas pela Rede PathoSpotter a testes de uso para coleta de informações sobre eventuais problemas;
- 2) Coletar informações que permitam melhor a qualidade das plataformas em teste.

**Avaliação dos Riscos e Benefícios:**

**Riscos:**

Cansaço durante a execução dos testes das ferramentas computacionais. Constrangimento por está sendo observado durante o uso.

**Benefícios:**

Oportunidade de aprendizado sobre desenvolvimentos em patologia computacional e contribuição com o desenvolvimento de sistemas que possibilitarão democratizar o acesso a exames anatomo-patológicos.

**Comentários e Considerações sobre a Pesquisa:**

A pesquisa apresentada é relevante e a equipe envolvida possui experiência para a condução da estudo e a execução metodologia proposta.

**Considerações sobre os Termos de apresentação obrigatória:**

Vide campo "Conclusões ou Pendências e Lista de Inadequações"

**Recomendações:**

Incluir no projeto como os resultados da pesquisa serão devolvidos/divulgados aos participantes da pesquisa.

**Conclusões ou Pendências e Lista de Inadequações:**

I. Análise da pendências apontadas no parecer consubstanciado deste protocolo emitido em 29 de Março de 2024:

1. "Incluir em "Riscos" no documento PB\_INFORMAÇÕES\_BÁSICAS\_DO\_PROJETO\_2283307 como os

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

**CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FIOCRUZ/BA**



Continuação do Parecer: 6.812.283

riscos aos participantes serão minimizados. Utilizar como referência as informações apresentadas no TCLE. RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, V". A informação foi incluída no projeto;

2. "Adequar o cronograma do estudo e se comprometer a iniciar o estudo somente após aprovação do protocolo pelo sistema CEP/Conep. Inserir a informação no projeto". Adequação realizada;
3. "Na folha de rosto, em patrocinador principal incluir "Financiamento Próprio". Esclarecer as fontes e a origem dos recursos". Informação corrigida;
4. "Anexar as cartas de anuênciade todos os participantes da equipe de pesquisa". Cartas anexadas;
5. "Incluir no projeto link para o currículo lattes dos participantes da equipe". Informações incluídas;
6. "Incluir os membros da equipe no formulário "Informações Básicas da Plataforma Brasil". Informações incluídas;
7. "Caso o projeto seja desenvolvido no IGM Fiocruz Bahia, anexar declaração de infraestrutura assinada pela diretora do instituto". Carta anexada;
8. "Incluir critérios de inclusão e exclusão. "Participantes: Indivíduos adultos de nível de graduação ou graduados nas diferentes áreas do saber. Estima-se que cerca de 1000 indivíduos serão arrolados nesses testes no período de 5 anos do projeto". RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, III". Informações incluídas;
9. "Descrever na metodologia o cenário de convocação dos participantes da pesquisa (onde, como e em que circunstância será apresentado o TCLE aos participantes). RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, III". Informação incluída;
10. "Apresentar o "Formulário sociodemográfico", o "Formulário de avaliação" e o roteiro da

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

**CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FIOCRUZ/BA**



Continuação do Parecer: 6.812.283

"Entrevista informal". RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, III". Instrumentos apresentados;

11. "Esclarecer quais tipos de imagens se refere a frase "..análise de imagens histológicas ciadas no projeto PathoSpo7er". RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, III". Esclarecimento realizado;

12. No TCLE. Resolução Nº 510, DE 07 DE ABRIL DE 2016, capítulo III seção II e RESOLUÇÃO Nº 466, DE 12 DE DEZEMBRO DE 2012, IV:

a) "Incluir no título do termo o grupo a quem se destina o documento. Por exemplo: "TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO DESTINADO PARA ADULTOS". Informação incluída;

b) "Incluir um cabeçalho no termo com as seguintes informações: nome do projeto, nome do pesquisador responsável e fonte de financiamento". Informações incluídas;

c) "Adicionar o número de contato telefônico e o endereço institucional de Bruna Neves de Oliveira, Dra. Emanuele Santos e Dr. Washington Santos". Informações incluídas;

d) "Informar que a participação é voluntária no parágrafo: "Sua participação é voluntária no referido estudo será no sentido de identificar problemas e dificuldades durante o uso da ferramenta proposta. O foco da pesquisa é o quanto fácil a qualidade da ferramenta e não visa conclusões sobre os usuários. O tempo estimado para a participação da pesquisa é em torno de 30 a 40 minutos. As etapas referentes à sua participação serão as seguintes". Informação inserida;

e) "Acrescentar um parágrafo tratando sobre como será assegurada a confidencialidade dos dados e divulgação dos resultados da pesquisa aos participantes". Parágrafo inserido;

f) "Adicionar de forma breve informações sobre as atribuições do CEP e da Conep, bem como os endereços e contatos telefônicos". Informações inseridas;

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

**CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FIOCRUZ/BA**



Continuação do Parecer: 6.812.283

- g) "Atualizar a numeração de páginas para o formato: ¿número de página/ número total de páginas¿.. Exemplo: 1/3; 2/3, etc". Informação atualizada;
- h) "Adicionar em todas as páginas, o campo para rubrica do (a) participante e pesquisador (a)". Informação adicionada;
- i) "Adicionar na última página do termo, o campo para assinatura do pesquisador (a)". Informação adicionada;
- k) "Acrescentar a informação que o termo será elaborado em duas vias originais, sendo que uma ficará com o (a) participante e a outra via com o pesquisador". Informações inseridas;
- j) "Adicionar informações que assegura que o (a) participante terá direito a indenização no caso de danos decorrentes do estudo". Informações adicionadas;
- l) "Adicionar previsão de resarcimento de despesas do (a) participante e compensação material (houver necessidade)". Informação adicionada.

**Considerações Finais a critério do CEP:**

O prazo para a execução desse projeto é 30/05/2029 conforme cronograma especificado nas informações básicas na plataforma.

Em cumprimento da Res. 466/2012 e Norma Complementar vigente, enviar para o CEP da instituição proponente relatório final em até um mês após o término da vigência do projeto conforme cronograma aprovado neste protocolo.

Relatórios parciais: a cada 06 meses;

Relatório final: 30/06/2029;

Caso haja inclusão de outra fonte de financiamento após esta aprovação, informar ao CEP como emenda ao protocolo incluindo o Termo de Outorga ou documento equivalente e realizando as alterações pertinentes na Folha de Rosto para indicar modificação do

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

**CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FIOCRUZ/BA**



Continuação do Parecer: 6.812.283

patrocinador principal.

The present study has been approved by the Comitê de Ética em Pesquisa do Centro de Pesquisas Gonçalo Moniz/FIOCRUZ (IORG0002090/OMB No. 0990-0279 valid until 01/27/2025).

The protocol and procedures presented in the project are in full accordance with the Brazilian legislation regarding the ethical standards in conducting research involving human beings (Res. CNS 466/2012).

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Projeto Detalhado / Brochura Investigador	PSPT_Plataforma_de_testes_do_Patho Spotter_projeto_MARCADO.pdf	07/05/2024 20:41:51	Jorge Clarêncio Souza Andrade	Aceito
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJECTO_2283307.pdf	29/04/2024 20:43:48		Aceito
Folha de Rosto	Folha_de_rosto.pdf	29/04/2024 20:43:21	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
Outros	RESPOSTA_AO_CEP_PLATAFORMAS .pdf	29/04/2024 20:41:28	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	PS_TERMO_DE_CONSENTIMENTO_LIVRE_E_ESCLARECIDO_DESTINADO_PARA_ADULTOS.pdf	27/04/2024 16:30:31	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
Outros	Entrevista_A.pdf	27/04/2024 16:24:35	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
Outros	Entrevista_B.pdf	27/04/2024 16:23:05	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
Declaração de concordância	Declaracao_diretoria_IGM.pdf	27/04/2024 16:16:06	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
Declaração de Pesquisadores	DECLARACAO_DE_ANUENCIA_Emanuele.pdf	27/04/2024 16:14:07	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
Declaração de Pesquisadores	DECLARACAO_DE_ANUENCIA_Bruna.pdf	27/04/2024 16:13:52	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br

CENTRO DE PESQUISAS  
GONÇALO MONIZ -  
FIOCRUZ/BA



Continuação do Parecer: 6.812.283

Projeto Detalhado / Brochura Investigador	PSPT_Plataforma_de_testes_do_Patho Spotter_projeto.pdf	05/02/2024 19:07:23	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_PSPT.pdf	05/02/2024 19:05:52	WASHINGTON LUIS CONRADO DOS SANTOS	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

SALVADOR, 08 de Maio de 2024

---

**Assinado por:**

**Jorge Clarêncio Souza Andrade**  
**(Coordenador(a))**

**Endereço:** Rua Waldemar Falcão, 121 na área de Escritórios do 3º Pavimento do Edifício Garagem, sala CEP

**Bairro:** Candeal

**CEP:** 40.296-710

**UF:** BA

**Município:** SALVADOR

**Telefone:** (71)3176-2285

**Fax:** (71)98814-7051

**E-mail:** cep.igm@fiocruz.br