

Estado da Arte sobre Engenharia de Requisitos e Explicabilidade em Sistemas Baseados em Aprendizado de Máquina

Lívia Mancine
Instituto de Informática
Universidade Federal de Goiás
Instituto Federal Goiano
Goiânia, Goiás, Brasil
liviamancine@discente.ufg.br

Taciana Novo Kudo
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás, Brasil
taciana@ufg.br

João Lucas Soares
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás, Brasil
soares.joao@discente.ufg.br

Renato F. Bulcão-Neto
Instituto de Informática
Universidade Federal de Goiás
Goiânia, Goiás, Brasil
rbulcao@ufg.br

ABSTRACT

With the recent growth in the use of Machine Learning (ML)-based software, concerns arise regarding explaining the results generated. Explanations help with transparency and increase stakeholder trust. Explainability, a term used to refer to these explanations, is considered a non-functional requirement (NFR) that substantially impacts the quality of ML systems. Explainability has become a mandatory requirement outlined in various laws in several countries. Additionally, Explainable Artificial Intelligence (XAI) is a field that studies methods supporting explainability in ML-based systems, focusing mainly on technical explanations. This study is not limited to technical explanations but provides a comprehensive overview of Requirements Engineering (RE) and the explainability requirement in AM-based systems. To achieve this, we planned and executed a Systematic Mapping Study protocol, adopting automatic searches in six databases. From the 200 returned articles, after applying selection criteria, we analyzed and reported the results of 27 articles. Our findings reveal that explainability is an emerging quality NFR in ML-based systems, challenging classical RE paradigms.

KEYWORDS

Engenharia de Requisitos, Explicabilidade, Aprendizado de Máquina, Mapeamento Sistemático da Literatura

1 INTRODUÇÃO

Sistemas multimídia e Web têm experimentado significativas transformações em seu desenvolvimento devido à crescente integração da Inteligência Artificial (IA), especialmente dos modelos de Aprendizado de Máquina (AM). Estes modelos estão cada vez mais presentes nesses sistemas, influenciando o conteúdo que consumimos diariamente. Um exemplo são os sistemas de recomendação Web que utilizam modelos de AM para identificar padrões nos dados

e fornecer sugestão aos usuários. Sistemas web populares, como Amazon¹ e Netflix², investem em pesquisas na área de recomendação e AM para melhorarem a experiência de seus usuários. Outro exemplo é o estudo de Souza *et al.* [16], que propõem a arquitetura de um sistema de recomendação conversacional que interage com o usuário e fornece explicações sobre as recomendações.

Essas explicações são baseadas em métodos da Inteligência Artificial Explicável (do acrônimo em inglês *Explainable Artificial Intelligence - XAI*). Estudos de métodos XAI geralmente visam fornecer explicações que atendam aos requisitos dos *stakeholders* internos, que são pessoas diretamente envolvidas no desenvolvimento, testes e implantação de sistemas de IA [1].

Neste sentido, o uso crescente de componentes de AM integrados a software levanta preocupações sobre a qualidade destes componentes, incluindo a confiabilidade dos resultados gerados pelos modelos de AM. Além disso, surgem questionamentos sobre a atribuição de responsabilidades em casos de resultados indesejáveis, bem como a compreensão sobre as razões pelas quais certos resultados são obtidos. Em muitos casos, é importante que os usuários finais e demais *stakeholders* entendam o porquê determinada decisão foi tomada pelos modelos de AM.

Nos últimos anos, tem sido amplamente relatado que os modelos de AM apoiam parcial ou totalmente a tomada de decisões indesejadas, que influenciam negativamente a vida das pessoas [30]. Essas decisões incluem injustiças e preconceitos, como o exemplo do concurso *Beauty.ai*³, o primeiro concurso internacional de beleza a usar algoritmos de AM como jurados. O concurso contou com seis mil participantes de cem países diferentes. Apesar da diversidade, a cor da pele foi um fator significativo na escolha dos vencedores; dos 44 vencedores, 36 eram brancos e apenas um preto.

Requisitos como o de explicabilidade têm se tornado uma preocupação emergente em sistemas baseados em AM. As explicações são recursos úteis para melhorar a transparência e a confiabilidade nesses sistemas. A explicabilidade refere-se a essas explicações e é considerada um requisito não funcional (RNF), que impacta

In: I Workshop de Revisões Sistemáticas de Literatura em Sistemas Multimídia e Web (WRSL+ 2024). Anais Estendidos do XXX Simpósio Brasileiro de Sistemas Multimídia e Web (WRSL+ 2024). Juiz de Fora/MG, Brazil. Porto Alegre: Brazilian Computer Society, 2024.

© 2024

ISSN 2596-1683

¹<https://aws.amazon.com/pt/pm/personalize/>

²<https://research.netflix.com/research-area/recommendations/>

³<https://www.theguardian.com/technology/2016/sep/08/artificial-intelligence-beauty-contest-doesnt-like-black-people>

substancialmente a qualidade dos sistemas de software [11, 13, 39]. Confiar decisões cruciais a um modelo de AM cria a necessidade de que esses modelos apresentem como característica importante a explicabilidade para o seu processo de tomada de decisão [23].

A capacidade de explicar está intimamente ligada à forma como os utilizadores finais e/ou especialistas de domínio podem compreender e explorar os resultados dos modelos de AM [30]. Assim, a explicabilidade, como um RNF importante, é considerada um requisito transversal. Isso significa que a explicabilidade é importante para todos os tipos de sistemas baseados em AM, incluindo sistemas multimídia e Web. A necessidade de explicabilidade transcende a simples funcionalidade do sistema. Portanto, é essencial estabelecer uma interação entre a Engenharia de Requisitos (ER) e a explicabilidade em sistemas baseados em AM, uma vez que a ER aborda de maneira sistemática os requisitos de software, incluindo requisitos funcionais e não funcionais.

Além disso, a explicabilidade em sistemas baseados em AM, está relacionada à confiança, justiça, análise de vieses e aspectos éticos, pois permite que os processos de tomada de decisão sejam compreensíveis e transparentes. Embora a interpretabilidade esteja associada aos modelos de AM, a explicabilidade sob a perspectiva da ER pode facilitar a compreensão das previsões complexas desses modelos, contribuindo para a confiança nos resultados gerados.

O objetivo deste Mapeamento Sistemático da Literatura (MSL) é investigar não apenas as explicações técnicas, mas fornecer um panorama da ER em relação ao requisito de explicabilidade em sistemas baseados em AM. Para isso, planejamos e executamos um protocolo para um MSL baseado na abordagem de [33].

A metodologia adotada incluiu uma análise sistemática das publicações disponíveis, permitindo-nos identificar lacunas na literatura e áreas que necessitam de maior investigação. Após a análise de 27 artigos relevantes para o tema, delimitou-se uma base sólida para futuras pesquisas e desenvolvimento de práticas que promovam explicabilidade em sistemas de AM. Embora exista uma preocupação da comunidade acadêmica em entender a relação da ER e sistemas baseados em AM [3, 40, 43], até o momento desta pesquisa, não encontramos estudos secundários como o mesmo objetivo que este.

Este artigo está assim organizado: a Seção 2 apresenta trabalhos relacionados; a Seção 3 descreve o planejamento do protocolo do MSL; a Seção 4 detalha a etapa de execução do MSL; a Seção 5 relaciona as respostas às questões de pesquisa, bem como sintetiza os resultados do MSL; por fim, a Seção 6 sumariza as conclusões e trabalhos futuros.

2 TRABALHOS RELACIONADOS

O processo da ER visa garantir que os requisitos devidamente elicitados, analisados, documentados e gerenciados atendam às necessidades dos *stakeholders*. No entanto, a natureza distinta dos processos de desenvolvimento de software tradicionais e aqueles baseados em IA gerou novas lacunas na ER para o desenvolvimento de sistemas baseados em AM. A engenharia de software tradicional envolve, basicamente, as atividades de ER, o projeto detalhado para implementação de um programa executável (principalmente a escrita de código) e tarefas de gerenciamento. Por outro lado, o desenvolvimento de um software baseado em AM inclui etapas

como a coleta de dados, a seleção de algoritmos de AM e o treinamento do modelo com base nos dados fornecidos, sendo a escrita de código uma atividade menos central no processo. Amershi *et al.* [5] compararam software tradicional com o software baseado em AM, destacando que: (a) descobrir, gerenciar e versionar dados para AM é mais complexo que em engenharia de software tradicional; (b) customizar e reutilizar modelos requer habilidades diferentes das encontradas em equipes de software; (c) componentes de IA e AM são mais difíceis de modularizar que os componentes de software convencionais. Nesse contexto, desenvolver software baseado em AM sem conhecimento detalhado, ou com conhecimento limitado do funcionamento interno do sistema, apresenta novos e significativos desafios para a ER.

Estudos anteriores investigaram abordagens de ER para sistemas de IA e AM, com ênfase em metodologias, ferramentas e técnicas específicas para esses sistemas [2, 40]. Ahmad *et al.* [2] conduziram uma MSL para identificar avaliações empíricas existentes, teorias emergentes e desafios relacionados à ER para IA. Eles derivaram uma lista de sinônimos para os termos “Engenharia de Requisitos” e “Inteligência Artificial” e personalizaram uma string de busca para cada base de dados, usando a seguinte sequência básica (*requirements engineering OR requirements process OR requirements elicitation OR requirements gathering OR requirements identification OR requirements analysis OR requirements validation OR requirements verification OR requirements specification OR “requirements development” OR “requirements documentation” OR “requirements management” OR requirements testing OR functional requirements OR requirements driven*) AND (*artificial intelligence OR machine learning OR expert systems OR deep learning OR computer vision OR natural language processing OR speech recognition OR machine intelligence OR AI OR ML OR chatbot OR expert systems OR self driving OR autonomous OR recommendation system OR robot*). A busca automática em seis bases, combinada com *snowballing* para frente e para trás, resultou na identificação de 43 estudos analisados. Eles destacaram linguagens de modelagem, como SysML, ontoML, diagrama de atividades e modelagem UML semi-formal, aplicadas nos estudos, além de enfatizarem aspectos éticos. Villamizar *et al.* [40] também conduziram uma MSL com o objetivo de delinear o estado da arte em ER para sistemas baseados em AM. A string de busca utilizada foi (*Software OR Applications OR Systems*) AND (*Machine Learning*) AND (*Requirements Engineering*). Utilizando uma estratégia de busca manual e *snowballing* para frente e para trás, os autores obtiveram 35 estudos. As principais descobertas indicaram que as etapas mais abordadas em ER foram elicitação e análise, com o uso frequente da técnica de *brainstorming*. Além disso, observou-se o emprego de métodos *ad hoc* para elicitar e garantir o cumprimento de RNFs. Os autores mencionaram que ainda não está claro quais ferramentas e técnicas tradicionais podem ser aplicadas da ER para AM. Além disso, notou-se que a maioria dos estudos sobre a intersecção entre ER e AM utiliza AM para fins da ER. Ambos os estudos, [2, 40], destacaram que aspectos de ética, confiança, justiça e explicabilidade são relevantes, mas são abordados apenas teoricamente e sem avaliação.

O estudo de Alves *et al.* [4] conduziu um *survey* para obter *insights* de profissionais sobre práticas e problemas enfrentados no ciclo de vida de requisitos em projetos de sistemas baseados em AM.

Foram coletadas 188 respostas de 25 países. Para análise, aplicaram-se métodos quantitativos, como *bootstrapping* com intervalos de confiança, e quantitativos, como codificação aberta e axial. Todos os participantes eram experientes em projetos de AM, com perfis predominantemente de cientistas de dados, seguidos por líderes de projetos, desenvolvedores, arquitetos de soluções e engenheiros de requisitos. A pesquisa indicou que líderes de projeto e cientistas de dados estão assumindo a responsabilidade pelas atividades de ER em sistemas baseados em AM, com notebooks sendo a principal ferramenta para documentação de requisitos. Os principais desafios a compreensão do problema e do negócio, a gestão de expectativas, requisitos pouco claros e a falta de disponibilidade e envolvimento de especialistas no domínio. No que refere aos RNFs, os profissionais destacaram preocupações específicas, como a qualidade dos dados, a confiabilidade do modelo e a explicabilidade do modelo, embora ainda falte aplicação prática sobre como lidar com esses requisitos.

Por outro lado, a pesquisa de Theis *et al.* [38] realizou uma análise estruturada da literatura e reportou os resultados de 48 artigos. Os autores relacionaram a explicabilidade com perspectivas técnicas e humanas no desenvolvimento de sistemas de IA, visando torná-los compreensíveis, aceitáveis e confiáveis, especialmente em sistemas críticos como tráfego aéreo. Os estudos analisados forneceram *insights* sobre o que é necessário para que as pessoas percebam uma IA como explicável, as informações necessárias para aceitação da IA e os métodos de representação e interação que promovem a confiança na IA. A pesquisa bibliográfica foi realizada de forma abrangente nas bases de dados Web of Science e Google Scholar, no repositório DLR eLib, no DLR Library Catalog, na NASA Technical Report Server, no Ebook Portal Central e no banco de dados das bibliotecas nacionais alemãs. Foram utilizados os seguintes termos de pesquisa (*explicabilidade OR rastreabilidade OR aceitação AND inteligência artificial AND raciocínio OR resolução de problemas OR representação do conhecimento OR planejamento automático OR programação automática OR percepção de máquina OR visão computacional OR robótica OR computação afetiva*). Os resultados indicam que os dois principais grupos de usuários são desenvolvedores, que necessitam de informações sobre as operações internas do modelo, e usuários finais, que precisam de informações sobre os resultados ou comportamento da IA. Além disso, os autores concluíram que a aceitação dos sistemas de IA depende de informações sobre as funções e o desempenho do sistema, de considerações éticas e de privacidade, bem como de informações adaptadas às preferências individuais para estabelecer confiança no sistema. Para atingir esses objetivos, é fundamental identificar as necessidades dos usuários e transformá-las em requisitos para o projeto do sistema de IA, constituindo uma etapa inicial para a ER. Esses requisitos devem ser validados e refinados para diferentes domínios de aplicação, servindo de base para as atividades de desenvolvimento.

Nosso estudo de MSL concentra-se particularmente no RNF de explicabilidade para sistemas baseados em AM. Exploramos aspectos relacionados à ER, incluindo a identificação de metodologias, ferramentas e notações de modelagem aplicadas a esses sistemas. Este enfoque busca uma compreensão abrangente das práticas atuais e a identificação de lacunas que possam orientar futuras pesquisas.

3 PROTOCOLO DE MAPEAMENTO SISTEMÁTICO

Para este estudo, utilizamos um protocolo de pesquisa que nos permitiu, por meio de MSL, averiguar os métodos, ferramentas, técnicas, abordagens e processo da ER para alcançar a explicabilidade em sistemas baseado em AM. O estudo seguiu as diretrizes propostas por Scannavino *et al.* [33]. Utilizamos a ferramenta *Parsif.al*⁴, uma ferramenta online de apoio à realização de revisões sistemáticas da literatura. Para mais detalhes sobre o protocolo e a extração dos dados, consulte o link disponível⁵.

A Figura 1 representa o MSL e inclui as fases de planejamento, condução e publicação. Primeiramente, um protocolo é planejado para que possa ser reproduzido posteriormente. Este protocolo inclui questões de pesquisa, *string* de busca, estratégia de pesquisa, fontes de estudos e critérios de seleção estudos e formulário de extração.

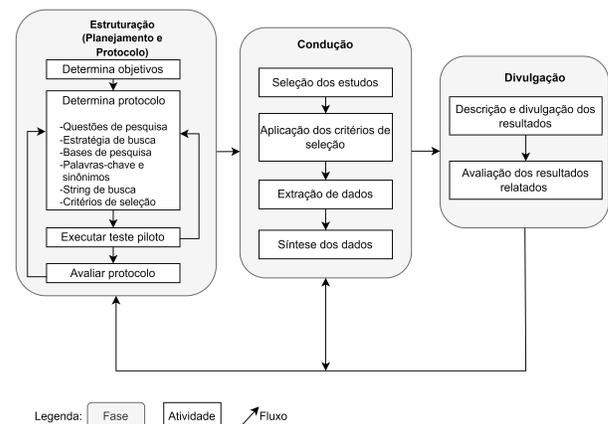


Figura 1: Fases e atividades do MSL.

Inicialmente, verificamos a necessidade de um MSL para este tema, dado que a literatura ainda é pouco explorada nessa área específica. Em seguida definimos as Questões de Pesquisa (QPs) fundamentais que guiaram nosso estudo, estabelecemos a *string* de busca para encontrar estudos relevantes, e determinamos critérios de inclusão e exclusão para seleção dos trabalhos.

3.1 Questões de Pesquisa

O principal objetivo desta pesquisa é *investigar o estado da arte de ER em relação ao requisito de explicabilidade em sistemas baseados em AM*. As questões de pesquisa formuladas derivaram deste objetivo, buscando esclarecer a proposta da ER para o requisito de explicabilidade, e quais os desafios e limitações enfrentados na integração deste requisito em sistema de AM. As seguintes questões de pesquisa (QPs) e seus objetivos foram definidas para este propósito:

⁴<https://parsif.al/>

⁵<https://doi.org/10.5281/zenodo.13227993>

QP1. Qual é o estado da arte da Engenharia de Requisitos (ER) em relação ao requisito de explicabilidade aplicado aos sistemas de AM?

O objetivo da QP1 é investigar frameworks, notações, linguagens de modelagem e ferramentas da ER para o desenvolvimento e definição de requisitos que garantam a explicabilidade em sistemas baseados em AM. Além disso, procurou-se identificar domínios mais preocupados em garantir a explicabilidade. Também foram investigados quais aspectos estão associados à explicabilidade e como os estudos têm conduzido suas pesquisas.

QP2. Quais são os desafios e lacunas apontados pela literatura da ER para sistemas de AM em relação ao requisito de explicabilidade?

O objetivo da QP2 é verificar os desafios e possíveis caminhos para pesquisas futuras sobre ER em relação ao requisito de explicabilidade em sistemas baseados em AM.

3.2 String de Busca

Para limitar a busca dentro do escopo da ER, explicabilidade e AM, elaboramos uma *string* de busca em três partes. A primeira parte incluiu termos relacionados ao AM e seus sinônimos, a segunda parte focou em ER e seus sinônimos, e a terceira parte abarcou o termo explicabilidade e seus sinônimos. Através de testes piloto, verificamos quais palavras-chave eram mais citadas nos estudos que tinham relação com os termos da nossa *string* de busca, conforme segue:

- Aprendizado de Máquina, notamos que palavras como: *machine learning*, *artificial intelligence*, *neural network*, *reinforcement learning* and *deep learning*, eram palavras que indicavam termos relacionados.
- Engenharia de Requisitos: os termos foram definidos de acordo com o SWEBOK⁶, tais como: *requirements engineering*, *requirements analysis*, *requirements elicitation*, *requirements management*, *requirements specification*, *requirements validation*, *requirements modeling*, *requirements process*.
- Explicabilidade: *explainability*, *XAI*, *interpretability*, *explainable*, *interpretable* and *trustworthy AI*, faziam referência ao termos explicabilidade.

Em seguida, realizamos diversas combinações em testes piloto utilizando as *strings* e suas respectivas partes. Este processo foi essencial para identificar a formulação mais eficaz da *string* de busca, que fornecesse evidências em relação ao nosso objetivo de pesquisa. Após várias iterações e refinamentos, definimos a *string* de busca final. Esta *string* foi aplicada aos títulos, resumos e palavra-chaves dos estudos selecionados, garantindo precisão na identificação de trabalhos. A *string* definida foi: *((machine learning OR artificial intelligence OR neural network) AND (requirements engineering OR requirements analysis OR requirements elicitation OR requirements*

management OR requirements specification OR requirements validation OR requirements modeling OR requirements process) AND (explainability OR XAI).

3.3 Estratégia de busca

A estratégia de busca adotada neste estudo foi a busca automática em bases bibliográficas. De maneira geral, para que um MSL forneça uma visão ampla de um dado tópico de pesquisa, é necessário realizar buscas automáticas em diferentes bases de dados relevantes [25]. Dada a natureza emergente e ainda pouco explorada no contexto da ER, acreditamos que a utilização de buscas automáticas nos permitirá explorar de forma eficaz a literatura disponível sobre ER e explicabilidade em sistemas de AM. Este método é particularmente adequado para identificar estudos primários, fornecendo uma base sólida para análise e síntese dos dados.

Selecionamos seis bases bibliográficas e motores de busca. De acordo com [25], essas bases podem ser úteis para obtenção de estudos no domínio de Engenharia de Software, e conseqüentemente da ER. São elas: *Engineering Village*, *IEEE Digital Library*, *Science Direct*, *Scopus*, *Springer Link* e *Web of Science*. Demais bases de dados, como a *ACM Digital Library*, não foram incluídas, pois, na data da realização deste estudo (março de 2024), não estavam disponíveis através do Portal de Periódicos da CAPES, que é a plataforma utilizada para acessar estudos científicos.

3.4 Critérios de seleção

Definimos a busca dos estudos publicados entre os anos de 2019 a 2024. Formulamos critérios de seleção que consistiram em critério de inclusão (CI) e critério de exclusão (CE). Os critérios de seleção, aplicados em todos os artigos que retornaram de acordo com a *string* de busca, são apresentados na Tabela 1.

Tabela 1: Critérios de Inclusão (CI) e Exclusão (CE) do MSL.

Critério	Descrição
CI	Estudos primários sobre ER que abordam explicabilidade em sistemas de AM
CE1	Não se trata de estudo primário
CE2	Não se trata de um artigo (por exemplo, prefácio ou resumo de periódicos ou anais de conferências)
CE3	O texto completo do estudo não está em inglês
CE4	Texto completo indisponível gratuitamente na Web ou na plataforma de periódico CAPES
CE5	É um versão preliminar ou resumida de outro estudo já incluso
CE6	A pesquisa não aborda Engenharia de Requisitos
CE7	A pesquisa não faz uma relação da ER para explicabilidade
CE8	O estudo não está no limite de publicação entre 2019-2024

3.5 Formulário de extração

A atividade de extração de dados requer um formulário cujos campos devem ser mapeados para as questões de pesquisa. Esses campos

⁶<https://www.computer.org/education/bodies-of-knowledge/software-engineering>

são preenchidos durante a leitura do texto completo de cada artigo. Elaboramos o formulário de extração de dados baseado nas QPs. A Tabela 2 apresenta o mapeamento entre os campos do formulário e as questões de pesquisa.

4 CONDUÇÃO DA SELEÇÃO DE ESTUDOS

Nesta seção descrevemos a etapa de seleção dos estudos primários. A primeira etapa consistiu na busca de artigos utilizando a string de busca, aplicadas aos títulos, resumos e palavras-chave nas bases bibliográficas e motores de busca selecionados em março de 2024, que retornou 200 artigos, como descreve a Tabela 3. Identificamos e removemos 70 artigos duplicados (do grupo de 200 estudos) com o apoio da ferramenta *Parsifal*.

Tabela 3: Número de artigos por fonte incluindo duplicados.

Base bibliográfica	Artigos retornados	Artigos aceitos
Engineering Village	34	0
IEEE Digital Library	33	06
Science Direct	2	0
Scopus	36	21
Springer Link	85	12
Web of Science	10	0
Total	200	39

Na segunda etapa, com a exclusão dos artigos duplicados, lemos o título, resumo e palavras-chave de cada um dos 130 artigos, sobre os quais aplicamos CI e CE e eliminamos 91 artigos (ver Tabela 4). Para decidir quais estudos seriam incluídos ou excluídos, aplicamos os CI e CE. Neste processo, dois pesquisadores revisavam os estudos de maneira independente e decidiam por incluir ou excluir o artigo. Em casos de incerteza, o artigo recebia uma anotação com a palavra “dúvida”. Nesses casos, um terceiro pesquisador, com maior experiência, realizava uma verificação adicional para determinar a inclusão ou a exclusão do artigo. Sempre que surgiam divergências nas decisões, os pesquisadores se reuniam para discutir e chegar a um consenso. Como resultado, selecionamos 39 artigos.

Durante a terceira etapa, no processo de extração de dados a partir da leitura completa dos artigos, quatro estudos foram eliminados pelo CE4, restando 35 estudos considerados como relevantes. Além disso, os pesquisadores identificaram que havia estudos que não se enquadravam no escopo definido. Esta discrepância ocorreu porque na etapa de seleção os artigos pareceriam indicar propostas adequadas à nossa pesquisa. Após relatar a situação ao pesquisador mais experiente, outros oito artigos foram eliminados pelo CE7, totalizando 103 estudos excluídos. Conforme descrito na Tabela 4, o critério CE6 excluiu a maior parte dos artigos. Isso significa que muitos estudos não focavam na ER para o requisito de explicabilidade em sistemas baseados em AM. Esse resultado sugere uma necessidade de maior clareza e detalhamento nas discussões sobre explicabilidade dentro do contexto da ER. Além disso, 32 artigos foram eliminados pelo CE6, pois não abordavam ER. Por último, 21 artigos foram excluídos por serem estudos secundários, mas que não tinham os mesmos objetivos deste estudo, conforme CE1, e um artigo pelo CE8, publicado em 2018.

No geral, extraímos informações de 27 artigos, estudos esses identificados ao longo desta pesquisa como E1 a E27 (ver Tabela 5). A Figura 2 retrata todo o processo de condução do MSL com o respectivo número de estudos primários escolhidos e removidos em cada atividade.

5 RESULTADOS E DISCUSSÃO

Esta seção apresenta a análise e síntese dos dados extraídos dos 27 estudos para responder às QPs do MSL.

5.1 Sobre a questão de pesquisa 1

QP1. Qual é o estado da arte da Engenharia de Requisitos (ER) em relação ao requisito de explicabilidade aplicado aos sistemas de AM?

A maioria dos artigos (20) é proveniente de conferências e *workshops* internacionais de grande relevância para a comunidade da ER, como *Requirements Engineering Conference Workshops (REW)*, *International Requirements Engineering Conference (RE)* e *International Conference on Software Engineering*. Isso sugere que o tema que relaciona ER, explicabilidade e AM, é de significativo interesse para a comunidade acadêmica. Os demais artigos (7), foram publicados em revistas da área de computação de modo geral. Além disso, a relação entre ER, explicabilidade e AM tem atraído grande atenção na comunidade científica. Nos últimos anos, esse tema ganhou destaque nas pesquisas (ver Figura 3).

Verificamos quais atividades da ER, incluindo elicitação, análise, especificação, validação e gerenciamento, eram abordadas pelos estudos analisados. Alguns desses estudos forneceram uma visão geral da ER aplicada à explicabilidade em sistemas baseados em AM e não contribuíram com técnicas específicas. Estes estudos incluíam documento de visão (de negócio), artigos de opinião e entrevistas tanto da indústria quanto da academia. Outros estudos contribuíram com ontologias, arquiteturas, metamodelos, modelos e *frameworks* concentrando-se mais nas primeiras atividades da ER, particularmente na elicitação e análise. A Figura 4 apresenta um gráfico que ilustra a quantidade de estudos focados em cada atividade ou grupo de atividades da ER.

Elicitação. A elicitação foi a atividade mais destacada, conforme demonstrado em: [6, 8, 9, 12, 24, 27, 34, 36, 41], correspondendo aos artigos E2, E6, E15, E26, E20, E12, E1, E17 e E13. Embora nem todos os estudos tenham apresentado técnicas de elicitação definidas, eles forneceram considerações importantes sobre as preocupações a serem abordadas ao elicitar requisitos de explicabilidade para sistemas baseados em AM.

Os estudos E1, E6, E15 e E26, identificaram a técnica de entrevista como uma possibilidade para elicitar requisito de explicabilidade. Isso se deve ao fato de que esses estudos utilizaram a avaliação empírica por meio de estudos de caso na indústria. Esses estudos sugerem que entrevistas diretas com especialistas do domínio podem ser uma abordagem eficaz para compreender os requisitos de explicabilidade. No trabalho E1 de Schellingerhout *et al.* [34], foram determinados os *stakeholders* que devem ser envolvidos no desenvolvimento de sistemas de IA, como Especialistas em Regulamentos e Engenheiros de IA, e suas relações. A explicabilidade foi apresentada como um RNF que impacta diretamente o escopo do projeto desses sistemas e deve ser uma preocupação desde o início do projeto, ou seja, durante a elicitação de requisitos. A pesquisa E6

Tabela 2: Informações do formulário de extração relacionado às QPs.

Informação	QP	Descrição
Metadados		Informações como o título do artigo, autores, data e veículo de publicação
Atividade da ER	1	Atividades da ER que o estudo aborda, como elicitación, análise, especificación, gerenciamento e validación
Técnicas e métodos associados a cada atividade da ER	1	Contribuição dos estudos em relação às técnicas e métodos
Domínio da aplicação	1	Informação sobre o cenário que existia uma preocupação da explicabilidade em sistemas baseados em AM
Tipos de requisitos abordados com a explicabilidade	1	Informação sobre quais requisitos estavam associados à explicabilidade
Stakeholders considerados no estudo em relação à explicabilidade	1	Codificado em: engenheiros de requisitos, especialistas de domínio, profissional de AM e/ou usuários finais
Tipo de contribuição	1	A contribuição foi codificado em: arquitetura, ferramenta, framework, metamodelo, modelo, método, métrica, ontologia, processo ou taxonomia
Estratégia de pesquisa	1	Classificação de estratégia empírica, de acordo com \cite{scannavino2017revisao}, incluindo estudos de casos, experimentos controlado, survey e simulação
Tipos de pesquisa	1	Classificação do tipo de pesquisa de acordo com \cite{scannavino2017revisao}, incluindo artigo de opinião, documento de visão, pesquisa de avaliação, pesquisa de validación, proposta de solução e relato de experiência
Limitação em relação a ER e a explicabilidade	2	Informações sobre as limitações em relação a ER e a explicabilidade em sistemas de AM
Trabalho Futuro	2	Tópicos de ER que apontam para trabalhos futuros

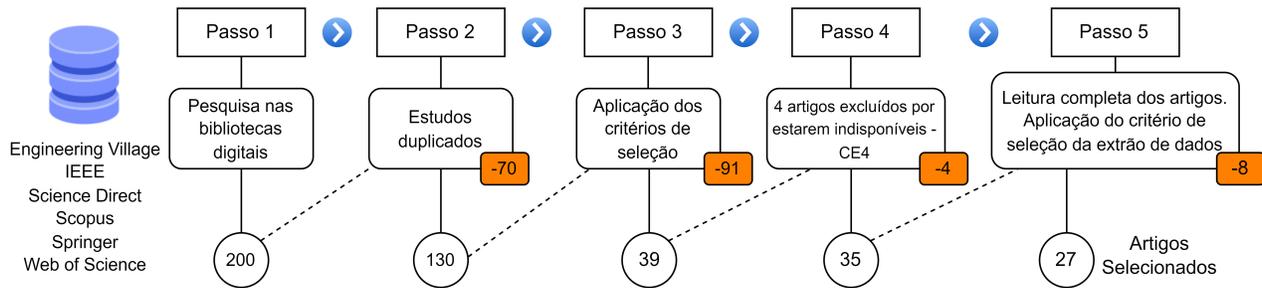


Figura 2: Visão detalhada da fase de condução: estudos primários selecionados e removidos em cada atividade.

Tabela 4: Número de artigos excluídos por critério de exclusão

	Leitura metadados	Extração de dados	Total
Critérios de exclusão			
CE1	21	0	21
CE2	0	0	0
CE3	0	0	0
CE4	0	4	4
CE5	0	0	0
CE6	32	0	32
CE7	37	8	45
CE8	1	0	1
Total	91	12	103

de Cabour *et al.* [8], propõe uma arquitetura que permite articular sistematicamente as explicações necessárias para os usuários finais em uma determinada tarefa, juntamente com os recursos XAI. Em E26 de Cirqueira *et al.*[12], utilizaram as ciências cognitivas para demonstrar um método de elicitación de requisitos baseado em cenários. No estudo E15 de Chazette *et al.* [9], foi desenvolvido um catálogo que abrange diversos aspectos de qualidade influenciados positiva e negativamente pela explicabilidade, o qual foi validado por especialistas.

Na pesquisa E2 de Aslam *et al.* [6], elaboraram uma proposta de solução através de Modelos Mentais para elicitare requisitos de explicabilidade e desenvolveram um Modelo Conceitual Orientado à Ontologia para facilitar o processo de aprendizagem dos usuários para uma melhor compreensão das explicações. Por outro lado, a pesquisa E13 de Vogelsang [41], propõe um framework para software autoexplicável, capaz de responder, em tempo de execução, a

Tabela 5: A relação dos 27 estudos analisados neste MSL.

ID	Estudo	Referência
E1	A Co-design Study for Multi-stakeholder Job Recommender System Explanations	[34]
E2	A Conceptual Model Framework for XAI Requirement Elicitation of Application Domain System	[6]
E3	A New Perspective on Evaluation Methods for Explainable Artificial Intelligence (XAI)	[37]
E4	A Requirements Engineering Perspective to AI-Based Systems Development: A Vision Paper	[18]
E5	An AI Chatbot for Explaining Deep Reinforcement Learning Decisions of Service-Oriented Systems	[29]
E6	An explanation space to align user studies with the technical development of Explainable AI	[8]
E7	An extension of iStar for Machine Learning requirements by following the PRISE methodology	[7]
E8	An ontology-based approach to engineering ethicality requirements	[19]
E9	Can Requirements Engineering Support Explainable Artificial Intelligence? Towards a User-Centric Approach for Explainability Requirements	[20]
E10	Dealing with Explainability Requirements for Machine Learning Systems	[28]
E11	Explainability as a Non-Functional Requirement	[26]
E12	Explainability Auditing for Intelligent Systems: A Rationale for Multi-Disciplinary Perspectives	[27]
E13	Explainable software systems	[41]
E14	Explainable software systems: from requirements analysis to system evaluation	[10]
E15	Exploring Explainability: A Definition, a Model, and a Knowledge Catalogue	[9]
E16	Holistic Explainability Requirements for End-to-End Machine Learning in IoT Cloud Systems	[31]
E17	How to Evaluate Explainability? - A Case for Three Criteria	[36]
E18	Human-centered XAI: Developing design patterns for explanations of clinical decision support systems	[35]
E19	Non-functional requirements for machine learning: understanding current use and challenges among practitioners	[21]
E20	On the Relation of Trust and Explainability: Why to Engineer for Trustworthiness	[24]
E21	Quality Characteristics of a Software Platform for Human-AI Teaming in Smart Manufacturing	[22]
E22	Quality-Driven Machine Learning-based Data Science Pipeline Realization: a software engineering approach	[15]
E23	Requirements Engineering for Explainable AI	[17]
E24	Requirements engineering for machine learning: Perspectives from data scientists	[42]
E25	Revisiting the Performance-Explainability Trade-Off in Explainable Artificial Intelligence (XAI)	[14]
E26	Scenario-Based Requirements Elicitation for User-Centric Explainable AI: A Case in Fraud Detection	[12]
E27	XAutoML: A Visual Analytics Tool for Understanding and Validating Automated Machine Learning	[44]

perguntas sobre seu comportamento passado, presente e futuro. Embora genérico, este estudo relaciona a ER na elaboração de técnicas para a explicabilidade no contexto da elicitaco.

Além disso, três dos estudos foram caracterizados como documentos de viso. Em [27], Langer *et al.*, estudo E12, apresentaram uma lista para prever como diferentes perspectivas podem se unir para garantir a explicabilidade de sistemas de AM. Para cada perspectiva, foram apresentados critrios que devem ser considerados para realizar uma auditoria. Este estudo fornece *insights* sobre a preocupaco em elicitaco requisitos tcnicos, psicolgicos, legais e ticos de diferentes *stakeholders* para garantir uma anlise abrangente da explicabilidade. Na pesquisa E17 de Speith [36], o autor motiva e defende trs critrios de qualidade da informaco que os sistemas devem fornecer para serem considerados explicveis: compreensibilidade, fidelidade e avaliabilidade. O autor examinou algumas abordagens XAI, como LIME, e avaliou as informaces que elas produzem com bases nos critrios estabelecidos. Em relao

ao LIME, os critrios de compreensibilidade e avaliabilidade indicaram uma avaliao positiva da explicabilidade. No entanto, o autor deixa claro que este estudo faz apenas um levantamento de questes tericas, uma vez que h pouco ou nenhum acordo sobre os mtodos de avaliao em explicabilidade. Nesse sentido, avaliando por uma perspectiva tcnica, este estudo permite entender qual abordagem XAI oferece os critrios citados para um determinado contexto, auxiliando na elicitaco do requisito de explicabilidade. Por ltimo, E20 de Kstner *et al.* [24] trs uma reflexo sobre a associao entre confiana e explicabilidade. Os autores argumentam que, para tornar um sistema baseado em AM confivel, garantir a explicabilidade pode ser extremamente til e, portanto, de grande importncia. Nesse sentido, deve-se avaliar como alcanar o critrio de confiana por meio da explicabilidade j no incio do projeto, na fase de elicitaco.

Anlise. Em relao a atividade de anlise da ER, o trabalho E5 de Metzger *et al.* [29], fornece uma arquitetura de chatbot (Chat4XAI)

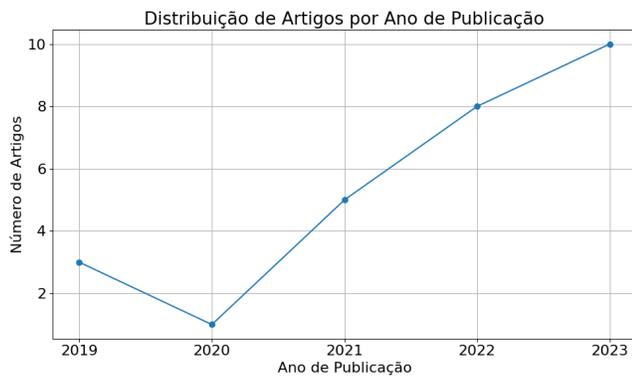


Figura 3: Distribuição de publicação entre os anos de 2019-2024 sobre ER, explicabilidade e AM.

que gera explicações em linguagem natural. Os autores se preocuparam com a Lei Regulamentar da IA da União Europeia (UE) e destacaram que a linguagem natural facilita uma melhor compreensão para usuários não técnicos. O Chat4XAI não exige a elicitación antecipada de requisitos e a classificação refinada das questões que os usuários possam fazer, pois aspectos mais específicos são tratados naturalmente por um grande modelo de linguagem (LLM) subjacente.

Embora os autores afirmem que a elicitación não seja uma exigência, os resultados demonstraram que o Chat4XAI com engenharia de prompt apresenta desempenho superior ao Chat4XAI com prompt *zero-shot* (que depende exclusivamente da capacidade do modelo de generalizar a partir de seu treinamento inicial). O estudo também comparou a fidelidade do Chat4XAI com a eficácia dos engenheiros de software, ou seja, o quão bem os engenheiros de software foram capazes de entender a tomada de decisão de um determinado modelo de AM comparado ao Chat4XAI. Os resultados demonstraram que Chat4XAI superou os engenheiros de software ao responder corretamente às perguntas sobre as explicações da tomada de decisão. Este estudo demonstra que, mesmo usando LLMs, é necessário ter um corpus relacionado ao contexto do domínio para se ter uma melhor eficácia nas explicações fornecidas. Ao avaliar o estudo, percebemos que um especialista em engenharia de requisitos deve participar do entendimento da necessidade de um corpus de domínio específico, o que enfatiza a importância de incorporar esse entendimento na atividade de análise de requisitos.

Em Li e Han [28], estudo E10, os autores propõem uma estrutura de análise de requisitos de explicabilidade em sistema de AM, utilizando modelos de objetivos contextuais para derivar métodos XAI de forma sistemática e automática. Os autores usaram a linguagem de modelagem *iStar*⁷ para modelar um framework de explicabilidade. Especificamente, pesquisaram e analisaram métodos XAI existentes com o intuito de recomendá-los a desenvolvedores de sistemas de AM. Este estudo utiliza uma abordagem baseada em objetivos para modelar requisitos de explicabilidade, focando no que as técnicas XAI visam oferecer. A investigação analisou a usabilidade da estrutura para desenvolvedores da pós graduação com vários níveis de conhecimento de explicabilidade. Neste estudo, a

explicabilidade focou exclusivamente em métodos XAI, abordando apenas os aspectos técnicos do projeto, como os desenvolvedores. Os demais *stakeholders*, que podem incluir usuários finais, especialistas de domínio e reguladores, não foram envolvidos no processo, o que limita a abrangência e a aplicabilidade das soluções propostas.

Elicitación e análise. Essas foram atividades integradas nos estudos [7, 10, 14, 19, 20, 26, 31, 44], correspondendo ao estudos E7, E14, E25, E8, E9, E11, E16 e E27. Estes estudos contribuíram com metamodelos, ontologias e frameworks.

No estudo E7 de Barrera *et al.* [7], os autores apresentaram uma extensão do *iStar* como um metamodelo para projetos de AM, permitindo identificar RNF, selecionar algoritmos mais adequados e analisar as fontes de dados disponíveis, alinhando-os aos objetivos do projeto. Eles incluíram um conjunto de diretrizes para facilitar sua aplicação, como um questionário que orienta especialistas em AM em entrevistas com especialistas do domínio. Além disso, forneceram um conjunto de tabelas que refletem como os modelos de AM existentes contribuem para RNFs e quais métricas de AM podem ser usadas para cada tipo de objetivo. Entre os RNF, destacaram a explicabilidade como importante nesse contexto. No entanto, o estudo não aborda todas as preocupações relacionadas à explicabilidade, concentrando em algumas tarefas de AM, como classificação, regressão e clustering, que são modelos mais clássicos de AM, e não incluíram áreas como *Deep Learning*. Além disso, os autores se preocuparam apenas com os métodos XAI que fornecem explicações técnicas, não avaliando aspectos mais amplos da explicabilidade que poderiam ser relevantes para diferentes *stakeholders*.

Na pesquisa E16 de Nguyen *et al.* [31], apresentaram uma abordagem holística para a cobertura dos requisitos de explicabilidade, a partir da definição dos *stakeholders*. Os autores mapearam as etapas para elicitación requisitos de explicabilidade e destacaram a necessidade de: i) identificar partes interessadas individuais concretas, ii) coletar os requisitos por meio de inquérito ou entrevista, e iii) continuar a atualizar os requisitos à medida que o desenvolvimento avança, por exemplo, a cada sprint se o método de desenvolvimento Ágil for empregado. Os autores ressaltam que ainda faltam técnicas e ferramentas para gerenciar a explicabilidade em contexto de sistemas baseados em AM, como plataforma de gerenciamento para acompanhar a documentação de requisitos, validar teste e entender como explicar questões relevantes em termos de conjunto de dados aos *stakeholders*. Tanto o estudo E7 quanto E16 realizaram uma pesquisa de avaliação com estudo de caso na indústria.

No artigo E8 de Guizzardi *et al.* [19], os autores apresentaram a Engenharia de Requisitos Baseada em Ontologia (ObRE) como um método para elicitación e analisar requisitos éticos, focando em dois importantes princípios de eticidade: explicabilidade e autonomia. A ontologia proposta foi validada para verificar se atende aos objetivos estabelecidos pela UE para o desenvolvimento de sistemas éticos. Utilizando um cenário de carro autônomo, os autores criaram tabelas de requisitos para garantir a conformidade com os princípios de explicabilidade e autonomia. Adicionalmente, empregaram um modelo de metas baseado no *iStar*, que retrata a dependência de cada um dos *stakeholders* e do carro autônomo. Embora o estudo demonstre uma preocupação significativa com requisitos éticos definidos pela UE, abordando a explicabilidade e propondo uma ontologia para elicitación e análise de requisitos de tais requisitos, os

⁷<https://http://istarwiki.org/>

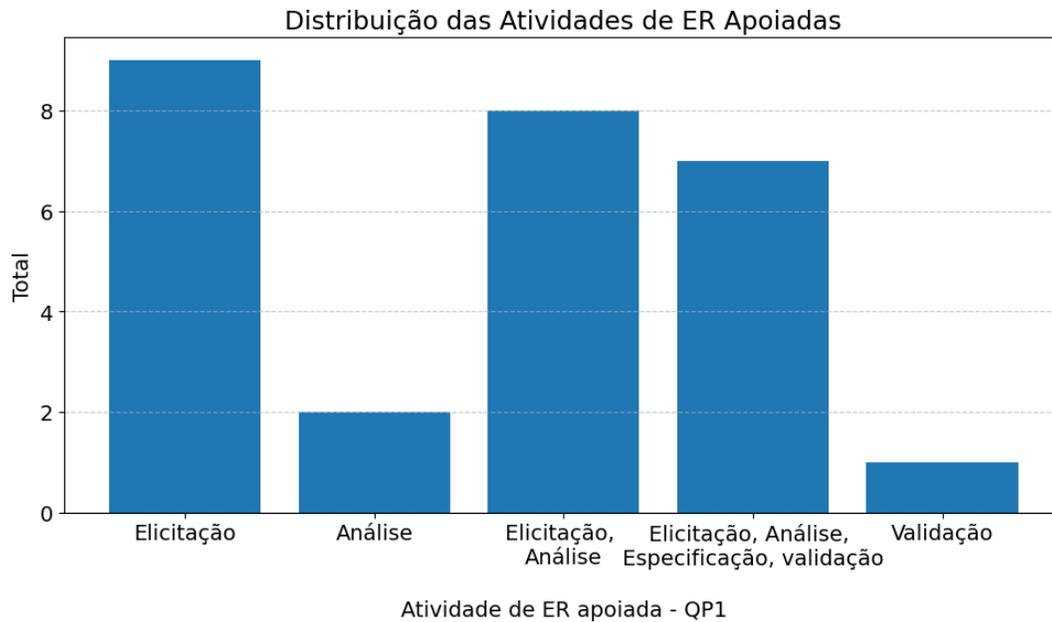


Figura 4: Atividades da ER apoiadas nos 27 artigos analisados neste MSL.

autores não implementou e nem validou a abordagem por meio de estudos de caso reais no domínio dos sistemas éticos, nem avaliou os resultados com especialistas.

Os autores Chazette *et al.*[10], estudo E14, realizaram uma pesquisa de validação por meio de estudo de caso acadêmico e desenvolveram quatro artefatos: uma definição de explicabilidade, um modelo conceitual, um catálogo de conhecimento e um modelo de referência para sistemas explicáveis. Esses artefatos visam apoiar engenheiros de software e de requisitos na compreensão da definição de explicabilidade e na sua interação com outros aspectos de qualidade no desenvolvimento de sistemas baseados em AM. No entanto, a correção dos modelos elaborados não foram avaliados e mais estudos são necessários para validar os modelos na prática. O estudo de E27 de Zöller *et al.*[44], enfatizou a importância do engenheiro de requisitos no desenvolvimento de sistemas baseados em AM, destacando a relação entre explicabilidade e desempenho. Os autores validaram uma ferramenta de análise visual intitulada *eXplainable Automated Machine Learning* (XAutoML), que visa compreender e validar procedimentos do AutoML (ferramenta que automatiza cada etapa do fluxo de trabalho de AM), tornando as otimizações do AutoML transparentes e explicando os modelos de AM produzidos.

O XAutoML permite a análise e compreensão de *pipelines* de AM, auxiliando profissionais de AM e da ER na avaliação da relação entre desempenho e explicabilidade. O estudo coletou e analisou requisitos de explicabilidade de uma base diversificada de usuários, incluindo 16 cientistas de dados, 11 especialistas de domínio e 9 pesquisadores de AutoML. Em relação à ER, a pesquisa revelou que, mesmo dentro de apenas um único grupo de usuários, como os especialistas de domínio, não há uma preferência clara sobre quais explicações devem estar disponíveis. Isso pode expor os usuários a

informações indesejadas. Por outro lado, os pesquisadores de AutoML solicitaram explicações de componentes internos específicos para depuração detalhada do sistema. Nesse sentido, um profissional da engenharia de requisitos poderia avaliar quais explicações seriam relevantes para determinados *stakeholders*. No entanto, a ferramenta precisa ser aprimorada para fornecer as interações comuns entre *stakeholders* e os sistemas AutoML em uma estrutura padronizada que acople análise visual otimizada. Uma limitação significativa é que o estudo não aborda completamente como balancear a necessidade de explicações de diferentes grupos de usuários, o que pode levar a dificuldades na implementação prática de soluções de explicabilidade.

E por fim, três estudos desenvolveram proposta de solução. Em [20], Habiba *et al.*, artigo E9 desenvolveram um framework com ênfase em uma abordagem centrada no usuário para requisitos de explicabilidade em AM. O estudo E11 de Köhl *et al.* [26], forneceram um catálogo baseado em diferentes noções de explicabilidade e requisitos de alto nível que as pessoas têm em mente quando exigem explicabilidade. Os autores enfatizaram a compreensão, utilizando uma abordagem multidisciplinar que inclui resultados da psicologia e das ciências cognitivas para avaliar se algo é realmente uma explicação e como as pessoas reagem a diferentes tipos de explicações. Os autores Crook *et al.* [14], do estudo E25 destacaram que os métodos XAI devem ser considerados pelos engenheiros de requisitos ao decidir sobre os modelos de AM na etapa de elicitação. Eles ressaltam que vários pontos devem ser considerados, como recursos de desenvolvimento, detalhes de domínio e riscos. Com base nisso, desenvolveram um framework que explora o compromisso entre desempenho, explicabilidade e tempo no contexto da IA a partir da perspectiva da ER. Essas propostas de solução exploram especialmente as atividades de elicitação e análise de requisitos de

explicabilidade. Embora apresentem soluções para explicabilidade no contexto de sistemas de AM, são apenas propostas de solução e carecem de experimentação em casos reais.

Elicitação, análise, especificação e validação. Estudos que trataram este conjunto de atividades da ER, apresentaram algumas características em comum. Os estudos E19 e E21 [21, 22] realizaram entrevistas para entender características e desafios da ER para sistemas baseado em AM. Em E19 validaram o estado da prática em relação aos RNFs de sistemas de AM, por meio de entrevistas com profissionais da academia e da indústria. Diversos RNFs foram identificados, incluindo a explicabilidade, que se destacou como uma preocupação nas principais atividades da ER. Como resultado, o grupo profissional destacou o RNF de desempenho, enquanto o grupo acadêmico enfatizou a explicabilidade. Na pesquisa E21, os autores entrevistaram profissionais da indústria para examinar a relevância de 11 características de qualidade para a formação de equipes humano-IA no desenvolvimento de sistemas de IA. As características de qualidade incluíam as 11 características do padrão ISO 25010:2011 para qualidade de software (SQuaRE) e 3 características de qualidade específicas de IA: confiabilidade, explicabilidade e auditabilidade. Neste estudo, a explicabilidade não foi um requisito destacado, sendo o desempenho uma preocupação mais proeminente. Estes estudos demonstram que a indústria ainda tem uma preocupação maior com o requisito de desempenho em comparação à explicabilidade. Isso sugere que os profissionais de desenvolvimento tendem a priorizar o desempenho, possivelmente devido à falta de profissionais de engenharia de requisitos que possam fornecer um balanço adequado entre explicabilidade e desempenho. A pesquisa E24 de Vogelsang e Borg [42], entrevistaram profissionais da indústria (cientistas de dados) e os resultados mostraram que esses profissionais tomam muitas decisões visando melhorar seus modelos de AM, focando no desempenho. Para essa tarefa, utilizam e referem-se a conceitos e medidas técnicas que muitas vezes não são bem compreendidos pelos clientes. Os autores concluíram que mudanças no paradigma de desenvolvimento, como no caso de sistemas baseados em AM, também exigem mudanças na ER. Assim, o desenvolvimento de sistemas de AM exige que os engenheiros de requisitos estejam cientes dos novos requisitos de qualidade, como a explicabilidade, e que extraíam esses requisitos do ponto de vista do usuário. Este estudo corrobora com os estudos E19 e E21, onde demonstram que os profissionais da indústria ainda têm uma grande preocupação com o desempenho dos modelos. Portanto, há uma demanda por engenheiros de requisitos que sejam capazes de basear suas decisões e conceitos técnicos em uma compreensão e análise completas das necessidades e do contexto dos clientes.

Em E18, Schoonderwoerd *et al.* [35], realizaram uma pesquisa de avaliação com um estudo de caso na indústria e descreveram uma abordagem de melhores práticas para *design* de um método XAI centrado no ser humano chamado DoReMi. Avaliaram a explicabilidade por meio de interfaces de usuário (IU) e conduziram um estudo de caso usando diagnósticos de um sistema de apoio à decisão clínica. Usaram entrevistas e avaliaram documentos para eliciar requisitos. Os autores estavam interessados em analisar a explicabilidade para especialistas de domínio, especificamente no campo médico. A abordagem DoReMi forneceu o primeiro conjunto de requisitos de usuário e padrões de design de IU para um

sistema explicável de apoio à decisão em saúde infantil. As avaliações com os médicos mostraram que eles realmente precisam de explicações sobre os resultados da IA, especialmente para ajudar a mitigar diagnósticos falsos positivos, evitando ao mesmo tempo falsos negativos. As IUs devem, portanto, ser projetadas de forma que as informações apresentadas sejam compreensíveis, ao mesmo tempo que apoiam o aprendizado sobre o sistema. Por exemplo, uma das IUs que apresentava a matriz de confusão mostrando a sensibilidade e a especificidade (métricas utilizadas para avaliar a qualidade do classificador) do modelo de AM para todos os casos ou uma quantidade selecionada de casos que receberam o mesmo diagnóstico, foi considerada como fornecendo a maior informação sobre a precisão do sistema. Este estudo focou basicamente nas explicações para especialistas de domínio e demonstrou que o conhecimento técnico é importante para esse grupo de usuário. Portanto, é essencial identificar quais interfaces podem aumentar a compreensão do sistema, garantindo que as explicações técnicas sejam eficazes e acessíveis.

Os demais estudos desse grupo de atividades da ER são predominantemente propostas de solução, focando em pesquisas que sugerem investigações e planejamentos futuros. A pesquisa E22 de d’Aloisio [15], apresenta uma proposta definindo um framework baseado em modelos que orienta os cientistas de dados no desenvolvimento de *pipelines* de AM garantindo requisitos de qualidade, como a explicabilidade. Nessa mesma perspectiva o estudo E23 de Umm-e-Habiba [17], propõe um modelo de processo de referência que serve como um guia para os profissionais abordarem os requisitos de explicabilidade relacionados a sistemas baseados em IA. Por fim, o estudo E4 de Franch *et al.* [18], que é um documento de visão, questiona o papel que a ER deve desempenhar no desenvolvimento de sistemas baseados em IA, destacando o escopo dos RNFs. Os autores argumentam que existem novos tipos de RNFs especificamente relacionados a sistemas baseados em IA, cuja relevância se destaca neste contexto, como confiança, ética e explicabilidade. Eles enfatizam que a ER se tornará a pedra angular que coordena todas as funções, atividades e artefatos envolvidos no desenvolvimento de sistemas baseados em IA.

Validação. O estudo E3 de Speith e Langer [37] apresenta métodos para avaliar abordagens de explicabilidade de acordo com os aspectos do processo XAI. Com base em modelos que explicam os principais processos XAI, a proposta é validar informações que são realmente explicativas, que facilitam a compreensão e satisfazem desejos sociais. Os autores argumentam que, para validar uma explicação, é necessário que ela apresente essas características. Destacam que, para melhorar a qualidade das informações explicativas, devem ser considerados critérios como compreensibilidade, fidelidade e avaliabilidade. Este estudo é um documento de visão e, portanto, como os demais estudos com essa característica, apresenta apenas uma proposta de solução sem implementações práticas ou validações empíricas.

Ainda sobre a QP1, apenas dez estudos realizaram pesquisa de avaliação com estratégia empírica, incluindo estudos de casos na indústria, *workshop* com especialistas e questionários aplicados a profissionais, incluindo E1, E6, E7, E15, E16, E18, E19, E21, E26 e E24. Outros 5 estudos, E5, E8, E10, E14 e E27, realizaram pesquisa de validação, usando como estratégia estudo de caso acadêmico,

simulação e prova de conceito. Os demais estudos, 12 num total, focaram em apresentar proposta de solução e documento de visão: E2, E3, E4, E9, E11, E12, E13, E17, E20, E22, E23 e E25. Além disso, os estudos relatam preocupação com diferentes perfis de *stakeholders* em relação a explicabilidade, isso inclui usuários finais, engenheiros de requisitos, profissionais de AM e especialistas de domínios.

Em relação ao domínio de aplicação dos estudos, muitos focaram suas propostas em um contexto geral, enquanto outros exploraram cenários hipotéticos. Embora não tenhamos encontrado um estudo específico sobre sistemas multimídia e web, entendemos que os componentes de AM podem ser integrados a diversos tipos de sistemas, uma vez que a explicabilidade é um requisito não funcional (RNF) essencial nesses contextos.

Por exemplo, no estudo E5 de Metzger *et al.* [29], a adaptação em sistemas orientados a serviços é discutida por meio de um caso específico de um sistema de comércio eletrônico adaptativo chamado SWIM, que simula uma loja virtual real. O SWIM permite a adaptação do sistema para maximizar uma função de utilidade específica diante de cargas de trabalho variáveis. Essa abordagem é utilizada como um exemplo concreto para demonstrar a aplicação do Chat4XAI na interpretação das decisões de algoritmos de aprendizado por reforço profundo.

A necessidade de explicabilidade em sistemas multimídia e web é crítica para garantir a confiança e a transparência das decisões tomadas por algoritmos de AM. Quando esses algoritmos são aplicados em contextos web, como recomendações de conteúdo ou personalização da experiência do usuário em sistemas de *e-commerce*, é fundamental que os *stakeholders* compreendam como e por que determinadas decisões são tomadas. Isso não apenas aumenta a confiança do usuário final, mas também facilita a manutenção e a melhoria contínua dos sistemas, garantindo que eles operem de maneira esperada e justa.

Como apresentado no estudos analisados, devemos considerar que os métodos XAI tem uma grande importância no contexto da explicabilidade. A Figura 5 ilustra os principais componentes da XAI, proposto por [13]. Basicamente, a XAI é formada por dois componentes principais: os modelos de AM e os métodos XAI. Na figura 5, é apresentado um modelo de AM de previsão, onde o modelo calcula as previsões com base nos dados de treinamento, enquanto o método XAI é responsável por gerar explicações do funcionamento interno e das previsões do modelo de AM [13]. Assim, a XAI incorpora dois resultados, previsões e explicações.

No entanto, alguns métodos de explicabilidade foram desenvolvidos especificamente para determinados tipos de dados, como imagens ou dados tabulares, enquanto outros não dependem do tipo de dados. Além disso, a XAI apresenta abordagens que focam em explicações relacionadas à interpretabilidade dos modelos de AM. Uma dessas abordagens é o *post-hoc*, que analisa e interpreta o processo de tomada de decisão de um modelo de AM treinado após ele ter feito previsões, fornecendo *insights* sobre como o modelo chegou aos seus resultados [32]. Por outro lado, as abordagens *ante-hoc* são chamados de explicabilidade intrínseca, modelos transparentes ou de caixa de vidro, e são modelos de AM inerentemente interpretáveis [32]. Esses métodos se concentram na incorporação de técnicas de interpretabilidade diretamente na arquitetura do modelo, como Árvores de Decisão.

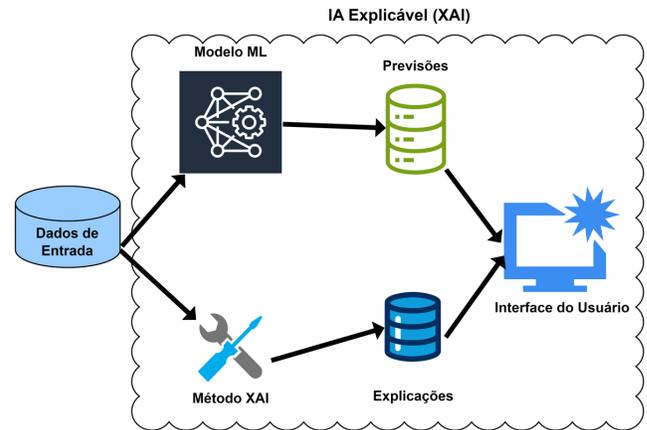


Figura 5: Principais componentes da Inteligência Artificial Explicável

Nesse sentido, é fundamental considerar a XAI e suas particularidades ao derivar requisitos de explicabilidade. Além disso, definir os *stakeholders* relevantes no processo de desenvolvimento de um sistema baseado em AM é crucial, pois diferentes stakeholders requerem diferentes tipos de explicações e formas de visualização. Por exemplo, um engenheiro de AM pode estar mais interessado em identificar as variáveis mais importantes para entender melhor o modelo. Por outro lado, um especialista de domínio no campo médico, pode requerer compreender o caminho percorrido pelo modelo de AM para chegar a um determinado diagnóstico, trazendo exemplos de outros casos semelhantes.

Portanto, entendemos que não é apenas importante definir o que explicar e a quem, mas também como explicar. Porém, a maioria das técnicas e ferramentas existentes da XAI são compreensíveis principalmente por *stakeholders* técnicos com experiência em AM, enquanto abordagens que atendam a todos os *stakeholders* relevantes são menos pesquisadas. Concluímos que os requisitos de explicabilidade devem ser definidos considerando os aspectos mencionados.

Por fim, o estado da arte da ER em relação ao requisito de explicabilidade aplicado aos sistemas de AM, revela que é uma área em crescimento. Nos últimos anos, especialmente em 2022 e 2023, a área ganhou destaque nas pesquisas. As atividades da ER mais exploradas foram a elicitación e análise, que são as fases iniciais do processo de ER. No entanto, a área carece de estudos empíricos, sendo a explicabilidade uma preocupação emergente em sistemas que integram AM. A explicabilidade é um RNF que pode contribuir com outros requisitos importantes em sistemas de AM, como conformidade legal, justiça e ética. De modo geral, a explicabilidade na perspectiva da ER deve ser mais explorada em todas as suas atividades, incluindo técnicas clássicas já consolidadas em software tradicionais, para validar se são suficientes ou precisam de adaptação neste contexto. A necessidade de entender como os sistemas de AM podem ser explicados de maneira clara e compreensível é urgente, mas a maioria dos estudos até agora tem se concentrado em abordagens teóricas sem validação em ambientes reais.

5.2 Sobre a questão de pesquisa 2

QP2. Quais são os desafios e lacunas apontados pela literatura da ER para sistemas de AM em relação ao requisito de explicabilidade?

Os desafios apresentados nos estudos estão principalmente relacionados à falta de pesquisas que reflitam a explicabilidade em ambientes reais. Dos 27 estudos analisados, 17 deles E2-E5, E8-E14, E17, E22, E23, E24, E25, e E27, conduziram pesquisas como documentos de visão, proposta de solução ou mesmo pesquisa de validação. Essa falta de experimentação prática constitui uma lacuna significativa apontada pelos estudos. Isso pode estar associado devido a crescente preocupação atual da explicabilidade em sistemas baseados em AM, que embora essencial em muitos contextos, ainda está em um estágio inicial de pesquisa. No entanto, a própria literatura sobre técnicas e abordagens de ER e explicabilidade em sistemas de AM, ainda não é suficiente. O estudo E2 de Aslam *et al.* [6], apresentou que a adoção de técnicas de modelagem conceitual apresenta desafios de pesquisa devido à literatura limitada sobre as abordagens interdisciplinares, além da falta de técnicas avançadas de modelagem essenciais para o seu desenvolvimento. Isso corrobora com a falta de estudos que apresentem estratégias empíricas na indústria.

A questão da generalização foi outro ponto abordado, pois os resultados obtidos em cenários específicos nem sempre podem ser aplicados a outros contextos. Por exemplo, no Brasil, é inaceitável pela lei⁸, práticas de discriminação de gênero ou raça. Para um sistema de AM utilizado em recrutamento de recursos humanos, essa informação não deverá ser relevante, porém, gênero e raça podem ser essenciais em um sistema de apoio à decisão baseado em AM para aplicações médicas. Nesse sentido, conforme o estudo E25, um engenheiro de requisitos deve identificar e excluir as características protegidas que não devem ser utilizadas pelo algoritmo de AM para discriminar amostras que dependem de cada contexto analisado [42].

Além disso, uma outra limitação significativa é a falta de engenheiro de requisitos especializados em sistemas de AM. De acordo com E9, o engenheiro de requisitos desempenham um papel crucial como mediadores entre os engenheiros de AM e os usuários finais [20], facilitando a comunicação e a compreensão mútua das necessidades e expectativas. Devido aos diversos *stakeholders* envolvidos no desenvolvimento de sistemas de AM, as explicações devem ser entregues de maneiras diferentes e adaptadas ao público-alvo.

Por exemplo, no cenário de recrutamento de recursos humanos, os desenvolvedor de AM necessitam de explicações técnicas detalhadas sobre como o modelo tomou suas decisões, permitindo-lhes analisar o modelo treinado e identificar características importantes nos dados de treino. Por outro lado, para os candidatos avaliados pelo sistema, eles precisam de explicações claras e transparentes, que forneçam informações compreensíveis sobre as razões por trás da decisão de seleção ou rejeição. Neste cenário, a presença de um engenheiro de requisitos é essencial para elicitar, analisar e especificar essas necessidades distintas, garantindo que as explicações fornecidas sejam adequadas e eficazes para cada grupo de *stakeholders*. O engenheiro de requisitos deve garantir que os requisitos de explicabilidade sejam bem definidos e atendam às exigências legais e éticas, bem como às expectativas dos usuários finais e dos

desenvolvedores de AM. Essa mediação é fundamental para assegurar a confiança e a aceitação dos sistemas de AM pelos diferentes *stakeholders* envolvidos.

De modo geral, a QP2 revela que existe um problema de generalização em relação aos contextos onde são empregados os sistemas baseados em AM. Isto é, os resultados e métodos desenvolvidos para um contexto específico muitas vezes não são diretamente aplicáveis a outros contextos, limitando a eficácia e a adaptabilidade das soluções de AM. Além disso, destaca-se a falta de engenheiros de requisitos especializados em AM, o que dificulta ainda mais a correta elicitação, análise e implementação dos requisitos de explicabilidade nesses sistemas. A ausência de profissionais qualificados para mediar entre as necessidades dos usuários finais e as capacidades técnicas dos sistemas de AM resulta em soluções que podem ser incompletas. Portanto, para avançar nesta área, é crucial investir na formação de engenheiros de requisitos com expertise em AM e desenvolver frameworks e metodologias que facilitem a generalização e adaptação de soluções de AM a diferentes contextos operacionais. Essas medidas são essenciais para garantir que os sistemas de AM possam ser amplamente aplicáveis, confiáveis e transparentes, atendendo às expectativas e necessidades de diversos *stakeholders*.

5.3 Síntese dos resultados

A Figura 6 apresenta o gráfico de bolhas que sintetiza as informações mais relevantes que extraímos e analisamos dos artigos aceitos neste MSL. Três eixos de informações compõem esses gráficos bolhas que mapeiam as fases da ER, o tipo de contribuição e os requisitos associados à explicabilidade. As atividades de elicitação e análise concentraram a maior parte das pesquisas, especialmente no que diz respeito a frameworks e modelos, que fornecem uma base estruturada para a especificação e análise de requisitos. No entanto, esses frameworks e modelos, bem como as demais técnicas avaliadas nos estudos, precisam ser mais explorados em ambientes reais. A ER oferece um arcabouço de técnicas, métodos e ferramentas para a elicitação, análise, especificação, validação e gerenciamento de requisitos. É necessário investigar se esse arcabouço é suficiente para abordar o RNF de explicabilidade em sistemas baseados em AM.

Além disso, os estudos demonstram que a explicabilidade pode contribuir significativamente para aspectos como confiabilidade, responsabilidade e transparência. Outros requisitos, como conformidade legal, justiça, satisfação e ética, também foram mencionados nos estudos analisados. A relação entre a explicabilidade e esses requisitos pode ser explorada para projetar a explicabilidade com base nesses fatores. Essas descobertas confirmam que a explicabilidade é essencial para melhorar a confiança e o entendimento entre o usuário e os sistemas baseados em AM.

6 CONCLUSÕES E TRABALHOS FUTUROS

De maneira geral, evidenciamos que a explicabilidade é um requisito de qualidade emergente em sistemas baseado em AM, desafiando os paradigmas clássicos da ER e levantando questões ainda sem resposta. Sistemas de software, incluindo multimídia e Web, estão cada vez mais integrando componentes de AM, impulsionados por demandas mercadológicas que buscam atender às necessidades e

⁸https://www.planalto.gov.br/ccivil_03/_Ato2023-2026/2023/Lei/L14611.htm

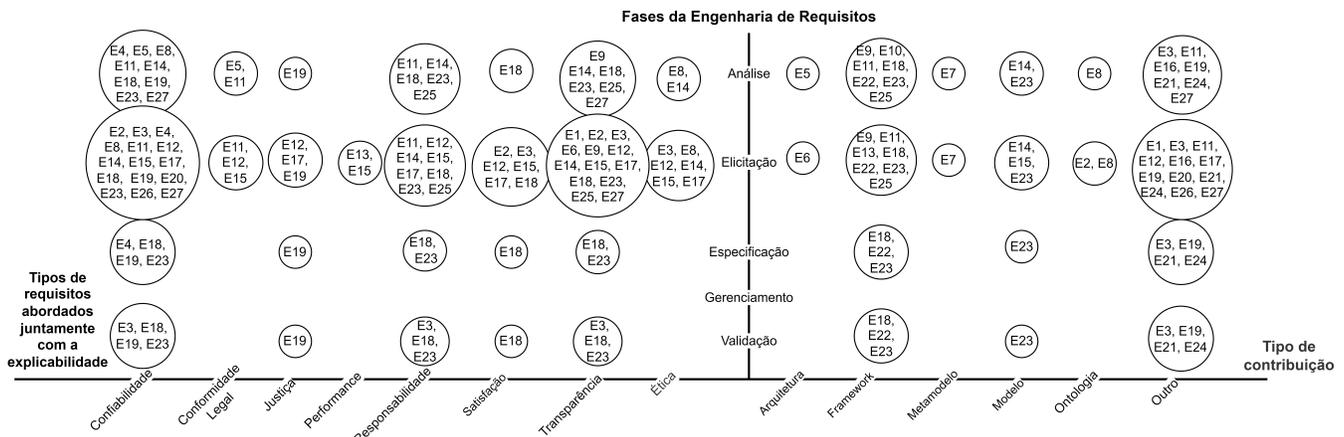


Figura 6: Gráfico de bolhas que mapeia as fases da ER, os requisitos associados à explicabilidade e o tipo de contribuição.

desejos dos seus clientes. Assim, a integração da explicabilidade em sistemas baseados em AM não é apenas uma questão técnica, abrangendo também questões éticas e sociais. Essa integração exige atenção contínua dos pesquisadores e desenvolvedores para mitigar os riscos associados e assegurar que a tecnologia beneficie a todos com equidade.

Com base nos principais resultados do nosso estudo, apresentamos uma síntese sobre os desafios e oportunidades para pesquisas futuras em ER voltadas para a explicabilidade em sistemas baseados em AM.

Atividades de ER. Nosso estudo revelou que as atividades de ER mais abordadas para projetar a explicabilidade em sistemas baseados em AM foram elicitación e análise. Isso sugere que a relação da ER e explicabilidade ainda está em processo de amadurecimento, com maior ênfase nas fases iniciais da ER. Por exemplo, vários estudos contribuíram com frameworks que fornecem uma visão geral para entender como elicitar componentes de AM, porém frequentemente focando na parte técnica da explicabilidade através de métodos XAI. Outros estudos abordaram soluções centradas no ser humano, mas como propostas de solução. Isto indica uma carência de estudos experimentais que validem essas abordagens em contextos reais. Portanto, há uma necessidade de mais pesquisa empírica para explorar a aplicabilidade prática dessas propostas, de modo a validar sua eficiência em ambientes reais e em diversos contextos de aplicação.

Aspectos relacionados à explicabilidade. Nosso estudo sintetizou os principais aspectos associados à explicabilidade, incluindo confiabilidade, responsabilidade, transparência e ética. A nossa análise sugere que, ao integrar a explicabilidade em sistemas baseados em AM, é importante avaliar esses aspectos subjacentes de maneira holística de acordo com o contexto.

Interdisciplinaridade. Trabalhar com explicabilidade exige uma abordagem interdisciplinar que considera as especificidades do domínio do problema, as necessidades dos *stakeholders* revelantes e implicações, como a ética e leis regulamentadoras. É essencial entender o contexto em que o sistema de AM opera para determinar quais informações devem ser explicadas e como essa explicação deve

ser estruturada. Além disso, a explicabilidade deve ser projetada de maneira a facilitar a auditabilidade e a verificação independente das decisões tomadas pelo sistema, reforçando a confiança dos usuários e a conformidade com regulamentos vigentes. Portanto, é importante que, nesse contexto, as particularidades da XAI sejam cuidadosamente consideradas para averiguar a contribuição nas explicações que devem ser fornecidas.

Avaliação. A avaliação da explicabilidade é um ponto crucial em sistemas baseados em AM. Nosso estudo, apontou a importância de desenvolver explicações que sejam compreensíveis, fiéis aos resultados gerados pelos modelos de AM e passíveis de avaliação pelos *stakeholders*. Isso garante que as informações explicativas fornecidas por esses sistemas sejam de alta qualidade. Esses critérios não apenas fornecem uma base estruturada para a avaliação da explicabilidade, mas também oferecem um caminho para o aprimoramento contínuo das capacidades explicativas dos sistemas de AM. Ao aplicar esses critérios de maneira sistemática, os desenvolvedores podem garantir que as explicações sejam claras, precisas e úteis para os usuários finais.

Este trabalho contribuiu para o campo de ER e AM ao apresentar os resultados de um MSL, e evidenciou importantes aspectos sobre a necessidade da ER para explicabilidade em sistemas baseados em AM, propondo avanços significativos na solução desse requisito. A explicabilidade é um RNF importante para sistemas que têm impactos relevantes, seja em um nível individual ou social, como por exemplo, sistemas de seleção de candidatos para emprego, avaliação de benefícios públicos, recomendação de conteúdo, e sistemas médicos. Tais sistemas podem integrar plataformas multimídia e Web, tornando relevante para a comunidade científica explorar a explicabilidade sob a perspectiva da ER.

Uma das limitações deste trabalho foi a falta de acesso em bases de dados como a *ACM Digital Library*, o que pode ter limitado a abrangência da pesquisa. Essa decisão foi necessária devido às restrições de acesso impostas pelas portal da CAPES. Futuras pesquisas poderiam considerar a inclusão dessas bases para uma análise mais completa e abrangente.

Além disso, não utilizamos outras estratégias de busca, como o *snowballing*. No entanto, dado que se trata de uma área de pesquisa emergente, a estratégia de busca automática pode ter sido suficiente para alcançar um número significativo de artigos relevantes, fornecendo evidências importantes para a área de pesquisa.

REFERÊNCIAS

- [1] Tedla Bayou Admekie and Sushmitha Pravin Karthick. 2024. An Exploration of Explainability for Internal Stakeholders: A Qualitative Study. (2024).
- [2] Khlood Ahmad, Mohamed Abdelrazek, Chetan Arora, Muneera Bano, and John Grundy. 2023. Requirements engineering for artificial intelligence systems: A systematic mapping study. *Information and Software Technology* 158 (2023), 107176.
- [3] Khlood Ahmad, Muneera Bano, Mohamed Abdelrazek, Chetan Arora, and John Grundy. 2021. What’s up with requirements engineering for artificial intelligence systems?. In *2021 IEEE 29th International Requirements Engineering Conference (RE)*. IEEE, 1–12.
- [4] Antonio Pedro Santos Alves, Marcos Kalinowski, Görkem Giray, Daniel Mendez, Niklas Lavesson, Kelly Azevedo, Hugo Villamizar, Tatiana Escovedo, Helio Lopes, Stefan Biffl, et al. 2023. Status quo and problems of requirements engineering for machine learning: Results from an international survey. In *International Conference on Product-Focused Software Process Improvement*. Springer, 159–174.
- [5] Saleema Amershi, Andrew Begel, Christian Bird, Robert DeLine, Harald Gall, Ece Kamar, Nachiappan Nagappan, Besmira Nushi, and Thomas Zimmermann. 2019. Software engineering for machine learning: A case study. In *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP)*. IEEE, 291–300.
- [6] Maria Aslam, Diana Segura-Velandia, and Yee Mey Goh. 2023. A conceptual model framework for XAI requirement elicitation of application domain system. *IEEE Access* (2023).
- [7] Jose M Barrera, Alejandro Reina-Reina, Ana Lavalle, Alejandro Maté, and Juan Trujillo. 2024. An extension of iStar for Machine Learning requirements by following the PRISE methodology. *Computer Standards & Interfaces* 88 (2024), 103806.
- [8] Garrick Cabour, Andrés Morales-Forero, Élise Ledoux, and Samuel Bassetto. 2023. An explanation space to align user studies with the technical development of Explainable AI. *AI & SOCIETY* 38, 2 (2023), 869–887.
- [9] Larissa Chazette, Wasja Brunotte, and Timo Speith. 2021. Exploring explainability: a definition, a model, and a knowledge catalogue. In *2021 IEEE 29th international requirements engineering conference (RE)*. IEEE, 197–208.
- [10] Larissa Chazette, Wasja Brunotte, and Timo Speith. 2022. Explainable software systems: from requirements analysis to system evaluation. *Requirements Engineering* 27, 4 (2022), 457–487.
- [11] Larissa Chazette and Kurt Schneider. 2020. Explainability as a non-functional requirement: challenges and recommendations. *Requirements Engineering* 25, 4 (2020), 493–514.
- [12] Douglas Cirqueira, Dietmar Nedbal, Markus Helfert, and Marija Bezbradica. 2020. Scenario-based requirements elicitation for user-centric explainable AI: A case in fraud detection. In *International cross-domain conference for machine learning and knowledge extraction*. Springer, 321–341.
- [13] Tobias Clement, Nils Kemmerzell, Mohamed Abdelaal, and Michael Amberg. 2023. Xair: A systematic metareview of explainable ai (xai) aligned to the software development process. *Machine Learning and Knowledge Extraction* 5, 1 (2023), 78–108.
- [14] Barnaby Crook, Maximilian Schlüter, and Timo Speith. 2023. Revisiting the performance-explainability trade-off in explainable artificial intelligence (XAI). In *2023 IEEE 31st International Requirements Engineering Conference Workshops (REW)*. IEEE, 316–324.
- [15] Giordano d’Aloisio. 2022. Quality-driven machine learning-based data science pipeline realization: a software engineering approach. In *Proceedings of the ACM/IEEE 44th International Conference on Software Engineering: Companion Proceedings*. 291–293.
- [16] Luan Soares de Souza, André Levi Zanon, Lucas Padilha Modesto de Araújo, and Marcelo Garcia Manzato. 2023. A multirun recommender system with explanations. In *Anais Estendidos do XXIX Simpósio Brasileiro de Sistemas Multimídia e Web*. SBC, 77–80.
- [17] Umm e Habiba. 2023. Requirements Engineering for Explainable AI.. In *RE*. 376–380.
- [18] Xavier Franch, Andreas Jedlitschka, and Silverio Martínez-Fernández. 2023. A requirements engineering perspective to AI-based systems development: A vision paper. In *International Working Conference on Requirements Engineering: Foundation for Software Quality*. Springer, 223–232.
- [19] Renata Guizzardi, Glenda Amaral, Giancarlo Guizzardi, and John Mylopoulos. 2023. An ontology-based approach to engineering ethicality requirements. *Software and Systems Modeling* 22, 6 (2023), 1897–1923.
- [20] Umm-E Habiba, Justus Bogner, and Stefan Wagner. 2022. Can requirements engineering support explainable artificial intelligence? Towards a user-centric approach for explainability requirements. In *2022 IEEE 30th International Requirements Engineering Conference Workshops (REW)*. IEEE, 162–165.
- [21] Khan Mohammad Habibullah, Gregory Gay, and Jennifer Horkoff. 2023. Non-functional requirements for machine learning: Understanding current use and challenges among practitioners. *Requirements Engineering* 28, 2 (2023), 283–316.
- [22] Philipp Haindl, Thomas Hoch, Javier Dominguez, Julien Aperribai, Nazim Kemal Ure, and Mehmet Tunçel. 2022. Quality characteristics of a software platform for human-ai teaming in smart manufacturing. In *International Conference on the Quality of Information and Communications Technology*. Springer, 3–17.
- [23] Vikas Hassija et al. 2023. Interpreting black-box models: a review on explainable artificial intelligence. *Cognitive Computation* (2023), 1–30.
- [24] Lena Kästner, Markus Langer, Veronika Lazar, Astrid Schomäcker, Timo Speith, and Sarah Sterz. 2021. On the relation of trust and explainability: Why to engineer for trustworthiness. In *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. IEEE, 169–175.
- [25] Staffs Keele et al. 2007. *Guidelines for performing systematic literature reviews in software engineering*. Technical Report. Technical report, ver. 2.3 ebse technical report. ebse.
- [26] Maximilian A Köhl, Kevin Baum, Markus Langer, Daniel Oster, Timo Speith, and Dimitri Bohlender. 2019. Explainability as a non-functional requirement. In *2019 IEEE 27th International Requirements Engineering Conference (RE)*. IEEE, 363–368.
- [27] Markus Langer, Kevin Baum, Kathrin Hartmann, Stefan Hessel, Timo Speith, and Jonas Wahl. 2021. Explainability auditing for intelligent systems: a rationale for multi-disciplinary perspectives. In *2021 IEEE 29th international requirements engineering conference workshops (REW)*. IEEE, 164–168.
- [28] Tong Li and Lu Han. 2023. Dealing with explainability requirements for machine learning systems. In *2023 IEEE 47th Annual Computers, Software, and Applications Conference (COMPSAC)*. IEEE, 1203–1208.
- [29] Andreas Metzger, Jone Bartel, and Jan Laufer. 2023. An AI Chatbot for Explaining Deep Reinforcement Learning Decisions of Service-Oriented Systems. In *International Conference on Service-Oriented Computing*. Springer, 323–338.
- [30] Diego Minatel, Nicolas Roque dos Santos, Angelo Cesar Mendes da Silva, Mariana Cúri, Ricardo Marcondes Maracini, and Alneu de Andrade Lopes. 2023. Fairness in machine learning for web systems applications. In *Proceedings of the 29th Brazilian Symposium on Multimedia and the Web*. 144–153.
- [31] My-Linh Nguyen, Thao Phung, Duong-Hai Ly, and Hong-Linh Truong. 2021. Holistic explainability requirements for end-to-end machine learning in IoT cloud systems. In *2021 IEEE 29th International Requirements Engineering Conference Workshops (REW)*. IEEE, 188–194.
- [32] Carl O Retzlaff, Alessa Angerschmid, Anna Saranti, David Schneeberger, Richard Roettger, Heimo Mueller, and Andreas Holzinger. 2024. Post-hoc vs ante-hoc explanations: xAI design guidelines for data scientists. *Cognitive Systems Research* 86 (2024), 101243.
- [33] Katia Romero Felizardo Scannavino, Elisa Yumi Nakagawa, Sandra Camargo Pinto Ferraz Fabbri, and Fabiano Cutigi Ferrari. 2017. Revisão Sistemática da Literatura em Engenharia de Software: teoria e prática. (2017).
- [34] Roan Schellingerhout, Francesco Barile, and Nava Tintarev. 2023. A Co-design Study for Multi-stakeholder Job Recommender System Explanations. In *World Conference on Explainable Artificial Intelligence*. Springer, 597–620.
- [35] Tjeerd AJ Schoonderwoerd, Wiard Jorritsma, Mark A Neerincx, and Karel Van Den Bosch. 2021. Human-centered XAI: Developing design patterns for explanations of clinical decision support systems. *International Journal of Human-Computer Studies* 154 (2021), 102684.
- [36] Timo Speith. 2022. How to evaluate explainability?—a case for three criteria. In *2022 IEEE 30th International Requirements Engineering Conference Workshops (REW)*. IEEE, 92–97.
- [37] Timo Speith and Markus Langer. 2023. A new perspective on evaluation methods for explainable artificial intelligence (xai). In *2023 IEEE 31st International Requirements Engineering Conference Workshops (REW)*. IEEE, 325–331.
- [38] Sabine Theis, Sophie Jentzsch, Fotini Deligiannaki, Charles Berro, Arne Peter Raulf, and Carmen Bruder. 2023. Requirements for explainability and acceptance of artificial intelligence in collaborative work. In *International Conference on Human-Computer Interaction*. Springer, 355–380.
- [39] Umm-E-Habiba. 2023. Requirements Engineering for Explainable AI. In *2023 IEEE 31st International Requirements Engineering Conference (RE)*. 376–380. <https://doi.org/10.1109/RE57278.2023.00058>
- [40] Hugo Villamizar, Tatiana Escovedo, and Marcos Kalinowski. 2021. Requirements engineering for machine learning: A systematic mapping study. In *2021 47th Euromicro Conference on Software Engineering and Advanced Applications (SEAA)*. IEEE, 29–36.
- [41] Andreas Vogelsang. 2019. Explainable software systems. *it-Information Technology* 61, 4 (2019), 193–196.
- [42] Andreas Vogelsang and Markus Borg. 2019. Requirements engineering for machine learning: Perspectives from data scientists. In *2019 IEEE 27th International Requirements Engineering Conference Workshops (REW)*. IEEE, 245–251.

- [43] Nobukazu Yoshioka, Jati H Husen, Hnin Thandar Tun, Zhenxiang Chen, Hironori Washizaki, and Yoshiaki Fukazawa. 2021. Landscape of requirements engineering for machine learning-based ai systems. In *2021 28th Asia-Pacific Software Engineering Conference Workshops (APSEC Workshops)*. IEEE, 5–8.
- [44] Marc-André Zöllner, Waldemar Titov, Thomas Schlegel, and Marco F Huber. 2023. Xautoml: A visual analytics tool for understanding and validating automated machine learning. *ACM Transactions on Interactive Intelligent Systems* 13, 4 (2023), 1–39.