



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

INTELIGÊNCIA ARTIFICIAL APLICADA À MANUTENÇÃO DE SOFTWARE: UMA REVISÃO DA LITERATURA

Priscila Bayer de Oliveira Simões

Escola Politécnica da Universidade de São Paulo (priscilabayer@usp.br)

Julieta Escorcia Escorcia

Escola Politécnica da Universidade de São Paulo (julieta.escorcia@usp.br)

Mauro de Mesquita Spinola

Escola Politécnica da Universidade de São Paulo (mauro.spinola@usp.br)

Rodrigo Franco Gonçalves

Escola Politécnica da Universidade de São Paulo (rofranco212@gmail.com)

Resumo

Além da manutenção tradicional, processos de manutenção de software podem se beneficiar da Inteligência Artificial. A exploração dos dados, gerados e coletados durante os processos, é uma ferramenta útil na descoberta de oportunidades de melhoria e acompanhamento durante a vida útil da aplicação. Foram identificados alguns campos da Inteligência Artificial aplicáveis no contexto: *Machine Learning*, *Deep Learning* e



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

Representation Learning. Este artigo tem como objetivo explorar a literatura relacionada aos temas “manutenção de software” e “inteligência artificial”, a fim de identificar o estado da arte das pesquisas e relações entre os conceitos. Através de uma revisão de literatura apoiada pela Web of Science e utilização de ferramentas de bibliometria e análise de conteúdo, verifica-se que é um tema explorado desde 2013, com o maior número de publicações em 2020. Verificou-se também que *machine learning* é a tecnologia com maior número de publicações e os processos de manutenção preditiva são os mais estudados. Pesquisas futuras podem estar relacionadas à evolução de software, um tema importante com menor número de publicações.

Palavras-chave: Inteligência artificial; Manutenção de software; Predição de falhas.

Abstract

Beyond traditional maintenance, software maintenance processes can benefit from Artificial Intelligence. The data, generated and collected during the procedures, and its exploitation bring benefits during the software's lifetime. Some fields of Artificial Intelligence have been identified were applicable: Machine Learning, Deep Learning, and Representation Learning. This paper aims to explore the literature related to the topics "software maintenance" and "artificial intelligence" in order to identify the state of the art of research and relationships between the concepts. Through a literature review supported by the tool Web of Science, the use of bibliometrics and, content analysis method, it was found that it is a topic explored since 2013, with the highest number of publications in 2020. It was also found that machine learning is the technology with the highest number of publications and predictive maintenance processes are the most studied publications. Future research may be related to software evolution, an essential topic with fewer publications.

Keywords: artificial intelligence; software maintenance; failure prediction.



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

Resumen

Además del mantenimiento tradicional, los procesos de mantenimiento de software pueden beneficiarse de la Inteligencia Artificial. La exploración de los datos generados y recogidos durante los procesos, es una herramienta útil en el descubrimiento de oportunidades de mejora y monitorización durante la vida útil del software. Se han identificado algunos campos de la Inteligencia Artificial aplicables en este contexto: *Machine Learning*, *Deep Learning* y *Representation Learning*. Este trabajo tiene como objetivo explorar la literatura relacionada con las áreas de "mantenimiento de software" e "inteligencia artificial", con el fin de identificar el estado del arte en la investigación, así como la relación entre los conceptos. A través de una revisión bibliográfica apoyada en la base de datos Web of Science, el uso de herramientas bibliométricas y un análisis de contenido, entre los hallazgos se encuentra que este es un tema explorado desde 2013, con un mayor número de publicaciones en 2020. Se encontró igualmente que *machine learning* es la tecnología con mayor número de publicaciones y los procesos de mantenimiento predictivo son los más estudiados. Futuras investigaciones pueden estar relacionadas con la evolución del software, un tema importante con un menor número de publicaciones.

1. INTRODUÇÃO

Ao observar o cotidiano e as referências no campo de Tecnologia da Informação, verifica-se que maior ênfase é dada ao projeto e implantação. Porém, a manutenção do software entregue ao final do ciclo é uma área de pouco prestígio e nem sempre recebe a atenção adequada (SOMMERVILLE, 2011). Todavia, os softwares precisam receber a manutenção adequada para que não haja prejuízos ao negócio devido a falhas ou indisponibilidades. A manutenção de software consiste no conjunto de atividades necessárias para mantê-lo operacional, de forma eficiente e com o menor custo possível



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

(BOURQUE; FAIRLEY, 2014). Caso a manutenção não receba a atenção necessária, pode se tornar uma atividade de alto custo, maior que em todo o ciclo de vida. Assim, investiu-se na construção de ferramentas para apoiar o desenvolvimento de software de qualidade, fácil manutenibilidade e que atenda aos requisitos. Isto pode se referir à manutenção eficiente; ou seja, no menor tempo, com o menor custo.

Existem softwares críticos, cujos requisitos de confiabilidade e disponibilidade são exigentes. Estes casos tornam crucial a identificação antecipada de possíveis pontos de falha, de forma a permitir que modificações sejam implantadas. Tais ações possuem como objetivo mitigar ou atenuar a ocorrência de falhas. Porém, o uso de software em todas as fases do negócio, desde as áreas operacionais até as estratégicas, faz com que a ocorrência de falhas em qualquer ponto do processo acarrete grandes problemas para a organização. Assim, entende-se como essencial a identificação antecipada de ações que evitem maus funcionamentos ou indisponibilidades.

A fim de exemplificar a oportunidade de pesquisa, executou-se uma busca simples de artigos. Inicialmente, foram encontradas mais de trezentas referências a respeito da aplicação de inteligência artificial em projeto e desenvolvimento de software. Porém, conforme apresentado na seção 3 – Metodologia, o número de publicações referentes à inteligência artificial aplicada à manutenção de software é expressivamente menor.

Este artigo tem como objetivo investigar, através de uma revisão da literatura, o estado da arte referente à aplicabilidade da Inteligência Artificial ao universo da manutenção de software. Para tal, são propostas as seguintes questões de pesquisa:

QP1: qual é a evolução da pesquisa ao longo dos anos, principais *journals* e autores referentes à utilização de Inteligência Artificial na manutenção de software?

QP2: quais são as aplicações de Inteligência Artificial utilizadas nas diferentes classificações de manutenção de software apresentadas na literatura?



Este trabalho contribui com o universo acadêmico através da verificação da extensão do conhecimento disponível na literatura referente à relevância do tema para pesquisas futuras. Nota-se que é um campo pouco explorado, o que pode indicar um novo território de pesquisa.

2. FUNDAMENTAÇÃO TEÓRICA

Esta seção apresenta as principais definições dos conceitos fundamentais explorados neste trabalho; ou seja, manutenção de software e inteligência artificial.

Manutenção de software

As definições e classificações referentes à manutenção de software presentes na literatura, tais como nos padrões de engenharia de software, são diversas. “Manutenção de software” é definida como processo de modificação para correção de falhas, melhoria de desempenho ou adaptação a um novo ambiente. São citadas as formas de manutenção corretiva e adaptativa (“IEEE Standard Glossary of Software Engineering Terminology”, [s.d.]). A manutenção de software mantém a capacidade de provisão de serviços através de monitoramento, registro de incidentes, tomada de ações corretivas, adaptativas e preventivas (“14764-2021 - ISO/IEC/IEEE International Standard - Software engineering - Software life cycle processes - Maintenance”, 2022). Outro padrão ISO/IEC/IEEE apresenta a construção de um plano de manutenção de software, que possui como objetivo a sustentação da capacidade de fornecimento adequado de serviços (“15026-4-2021 - ISO/IEC/IEEE International Standard - Systems and software engineering--Systems and software assurance -- Part 4: Assurance in the life cycle”, 2021). Em suma, os padrões ISO/IEC/IEEE referentes à manutenção de software apresentam, em comum, classificações referentes à correção de falhas, adaptação e aperfeiçoamento do software e prevenção de falhas.



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

Há novas definições para as modalidades de manutenção descritas anteriormente (“Systems and software engineering - Vocabulary”, 2010):

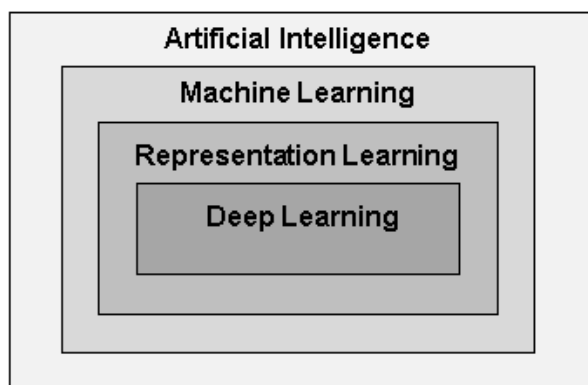
- Manutenção adaptativa: provê melhorias necessárias para a acomodação de mudanças no ambiente;
- Manutenção corretiva: modificação reativa aplicada a fim de corrigir falhas e problemas. Podem ser aplicados planos de ação emergenciais;
- Manutenção de aperfeiçoamento: aplicada na correção de erros latentes e melhorias sugeridas ou solicitadas.

Nos supracitados padrões, não foram identificados itens referentes à manutenção preditiva. Porém, de acordo com a análise dos artigos encontrados na execução da chave de busca descrita na seção 3, serão adotadas as classificações corretiva – evolutiva – preditiva com o intuito de explorar esta área.

Inteligência artificial

O termo “Inteligência Artificial” foi cunhado na década de 1950 com abordagens racionais e centradas no ser humano. As abordagens racionais são uma combinação de Engenharia e Matemática, enquanto as centradas no ser humano são parte das ciências empíricas. A fim de serem eficientes, as técnicas de Inteligência Artificial deveriam possuir a capacidade de extrair padrões de dados e criar seu próprio conhecimento, o que é conhecido como “*machine learning*”. *Deep learning*, um desdobramento do *representation learning*, permite que o aprendizado ocorra através de diversos níveis de abstração, que formam modelos de camadas de multiprocessamento (BORGES et al., 2021). A fim de ilustrar este conceito, os autores apresentam o diagrama da Figura 1:

Figura 1 – hierarquia entre as camadas da Inteligência Artificial.



Fonte: Adaptado de (BORGES et al., 2021)

3. MÉTODO

Foram selecionadas, para a elaboração deste trabalho, os métodos “bibliometria” e “análise de conteúdo” através de codificação. A bibliometria é um método que utiliza os dados bibliográficos de bases de publicações científicas a fim de construir representações gráficas relacionadas aos campos de estudo, desempenho de pesquisa e validação da estrutura do conhecimento (ZUPIC; ČATER, 2015). A codificação é utilizada para executar análises quantitativas dos principais construtos identificados durante a pesquisa.

Amostragem

O primeiro passo da amostragem é composto pela definição das bases de dados a serem consultadas, seleção dos termos que compõem a chave de busca e os tipos de documentos considerados no estudo. Definiram-se também os critérios de inclusão e exclusão de documentos. Por fim, será apresentada uma análise referente aos resultados deste estudo.

A base de publicações escolhida para exploração é a *Web of Science* por ser reconhecidamente confiável, amplamente utilizada na pesquisa científica e conter *journals* conceituados e utilizados em estudos de revisão de literatura.

Os critérios utilizados na seleção inicial dos documentos foram os de tipo “*article*”, “*early access*”, “*review*” e “*article/proceedings*”, em inglês. Não foi estabelecido limite temporal nas buscas; ou seja, publicações recentes e antigas foram levadas em consideração, sem a aplicação de filtro por ano. A primeira busca foi executada em 20 de julho de 2022, passou por algumas revisões e a última ocorreu em 06 de agosto de 2022. A chave de busca empregada é a seguinte: “*artificial intelligence*” AND (“*software maintenance*” OR “*software evolut**” OR “*Software Defect Prediction*” OR “*software unava**” OR “*system reliability*” or “*software failure*” or “*bug prediction*”). A busca pelos termos na base foi feita levando-se em conta o item “Tópico”. O operador lógico “and” foi utilizado para combinar os grupos de termos de busca; o operador “or” foi aplicado entre sinônimos a fim de retornar todos os documentos relacionados. O resultado desta busca é apresentado na Tabela 1.

Feita a busca, os resultados foram exportados e gerou-se um arquivo de metadados em formato BiBTEX. Este arquivo serviu como entrada para o processo de limpeza e consolidação para a formação posterior de uma base final que apoiará todo o trabalho. Este processo foi executado através de um script de tratamento de dados na ferramenta R Studio. Realizou-se uma verificação manual a fim de validar o arquivo de saída deste tratamento e a qualidade do processamento descrito. Por fim, aplicaram-se os critérios de seleção definidos a fim de gerar a base de dados final com as publicações selecionadas para leitura e estudo.

Tabela 1 – Pesquisa bibliográfica

Base	Chave de busca	Campos	Filtros	Quantidade
Web of Science	“ <i>artificial intelligence</i> ” and (“ <i>software maintenance</i> ” or “ <i>software evolut*</i> ” or “ <i>software defect prediction</i> ” or “ <i>software unava*</i> ” or “ <i>system reliability</i> ” or	Tópico	Inglês; “ <i>article</i> ”, “ <i>early access</i> ”, “ <i>review</i> ” e “ <i>article/proceedings</i> ”	133

	<i>“software failure” or “bug prediction”</i>			
--	---	--	--	--

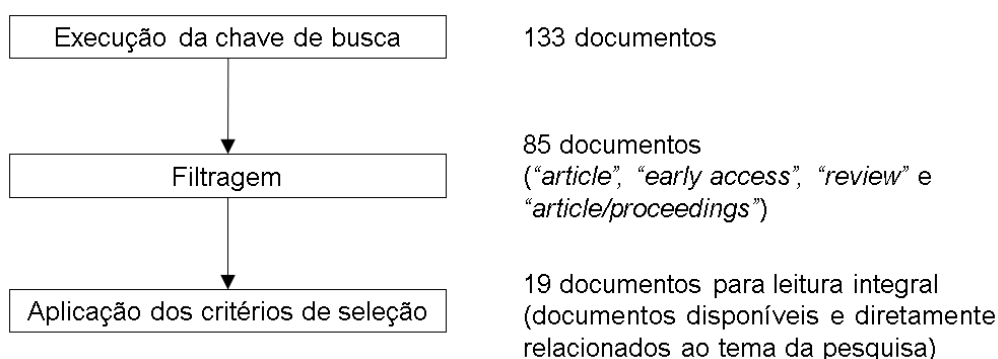
Fonte: autoria própria.

Os critérios de seleção foram definidos através da seguinte estrutura:

- Seleção por tipo de documento (artigos, acesso antecipado, revisão e *“article/proceedings”*, em inglês);
- Execução de *script* para processamento da base de metadados via ferramenta *R Studio*;
- Leitura dos títulos e resumos com o objetivo de selecionar apenas as publicações relacionadas ao tema do trabalho. Os demais foram desconsiderados e excluídos da seleção para a base final.

A Figura 2 ilustra a execução da sequência de etapas descritas anteriormente e os resultados obtidos em cada uma delas.

Figura 2 – Processo de amostragem



Fonte: autoria própria

Em suma, o processo de amostragem consistiu na consulta da chave de busca na base *Web of Science*, o que trouxe 133 documentos. Após a aplicação de filtros baseados por tipo de publicação, 85 documentos foram selecionados para análise dos títulos e resumos. Após a aplicação dos critérios de inclusão e exclusão, 19 artigos foram lidos na íntegra para o desenvolvimento deste trabalho.

Análise dos dados

A revisão de literatura pode ser apoiada por diferentes abordagens, tais como bibliometria, meta-análise e análise de conteúdo (CARVALHO; FLEURY; LOPES, 2013). Para a escrita deste artigo, foram utilizados os métodos “bibliometria” e “análise de conteúdo”. A interface web *Biblioshiny* foi empregada no desenvolvimento da bibliometria. Os metadados extraídos da base *Web of Science* foram utilizados para realizar as análises posteriores referentes à bibliometria (SECINARO et al., 2020). Em seguida foi construída a análise de conteúdo através da catalogação e leitura dos artigos selecionados. A Tabela 2 apresenta a aplicação dos métodos selecionados (bibliometria e análise de conteúdo) na resposta às questões de pesquisa.

Tabela 2 – questões de pesquisa, métodos e ferramentas

ID	Questão de pesquisa	Método	Ferramenta
1	Qual é a evolução da pesquisa ao longo dos anos, principais <i>journals</i> e autores referentes à utilização de Inteligência Artificial na manutenção de software?	Bibliometria	Biblioshiny
2	Quais são as aplicações de Inteligência Artificial utilizadas nas diferentes classificações de manutenção de software?	Análise de conteúdo; codificação	-

Fonte: autoria própria



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

Bibliometria

A bibliometria é um conjunto de métodos para extração de medições em um conjunto de dados bibliográficos. Diversas áreas de pesquisa utilizam métodos bibliográficos a fim de explorar impactos de área de pesquisa, grupo de pesquisadores ou até mesmo um artigo específico (COBO et al., 2011).

Neste trabalho foram construídos elementos gráficos utilizando-se a ferramenta *Biblioshiny*, além de uma análise exploratória no Excel. As análises desenvolvidas no *Biblioshiny* ilustram o impacto da produção por autor através do número de citações, mapa temático para verificar a relevância do tema, relação entre países, autores e palavras-chave. A análise exploratória demonstra o número de publicações por ano.

Análise de conteúdo

A análise de conteúdo é uma metodologia utilizada para explorar as referências por meio da quantificação do significado da língua falada ou escrita. Tal método fornece uma visão integrativa do texto e seus contextos relacionados para que o pesquisador compreenda o fenômeno de forma subjetiva (RENZ; CARRINGTON; BADGER, 2018).

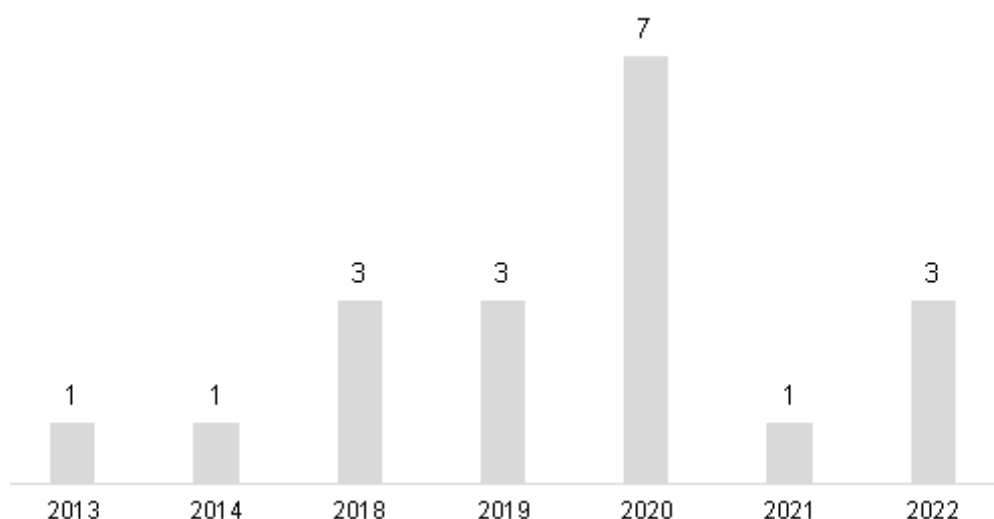
Tal análise segue a metodologia previamente descrita, a partir da codificação dos principais conceitos identificados em cada um dos artigos, com a finalidade de fornecer base e direcionamento para a posterior apresentação dos resultados do estudo. A fim de compor a análise final, foi feita uma contagem de ocorrências de cada um dos itens nos artigos, classificadas por grupos. As discussões apresentadas posteriormente são fundamentadas por tal processo de análise.

4. RESULTADOS

Esta seção apresenta os resultados das análises construídas através da bibliometria, de acordo com os procedimentos indicados em Metodologia. Construiu-se uma análise

exploratória com base na amostra bibliográfica extraída a fim de verificar e quantificar a produção científica. A Figura 3 apresenta a quantidade de publicações relacionadas ao tema por ano. Nota-se que é um campo de pesquisa já explorado nas últimas décadas, mas que recebeu maior atenção em 2020. Porém, o número de publicações demonstra uma possível oportunidade de pesquisa em um campo ainda não suficientemente explorado.

Figura 3 – publicações por ano



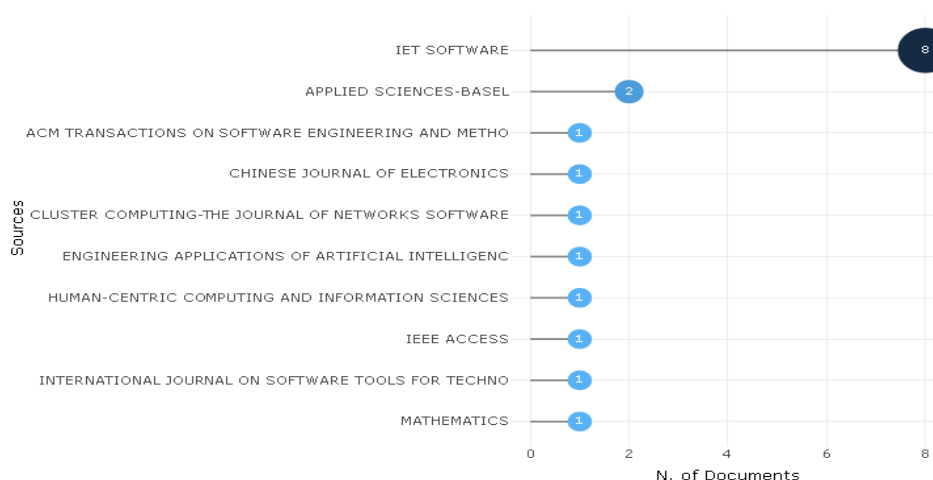
Fonte: autoria própria

Diversas áreas exploraram o assunto; porém, de forma esparsa. O *journal* com o maior número de publicações é a *IET Software*, com oito publicações, seguida por *Applied Sciences-Basel*, com duas publicações. As demais disponibilizaram apenas um documento referente ao tema da pesquisa. *IET Software* é um *journal* que publica pesquisas originais a respeito de todo o ciclo de vida do software (“*IET Software Impact Factor & Key Scientometrics*”, [s.d.]). *Applied Sciences-Basel* publica pesquisas

relevantes referentes a diversos temas, incluindo Engenharias, Sustentabilidade e Ciência dos Materiais (“About Applied Sciences”, [s.d.]). Ambos os *journals* são *golden open access*. Não é possível identificar uma clara concentração de publicações em áreas diversas. A Figura 4 apresenta uma síntese desta análise.

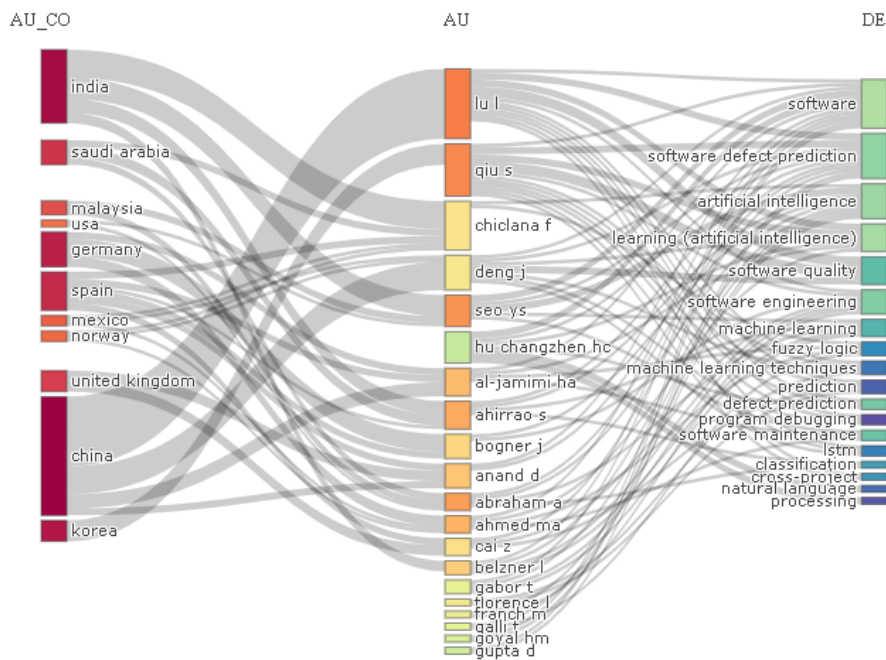
A Figura 5 representa, através de um gráfico do tipo *three fields plot*, o conjunto de relações entre países, autores e palavras-chave. Foram encontradas dezoito palavras-chave, vinte autores e onze países. O termo com o maior número de ligações é “software”, com o total de sete. O maior número de publicações foi escrito na China.

Figura 4 – publicações por *journal*



Fonte: autoria própria

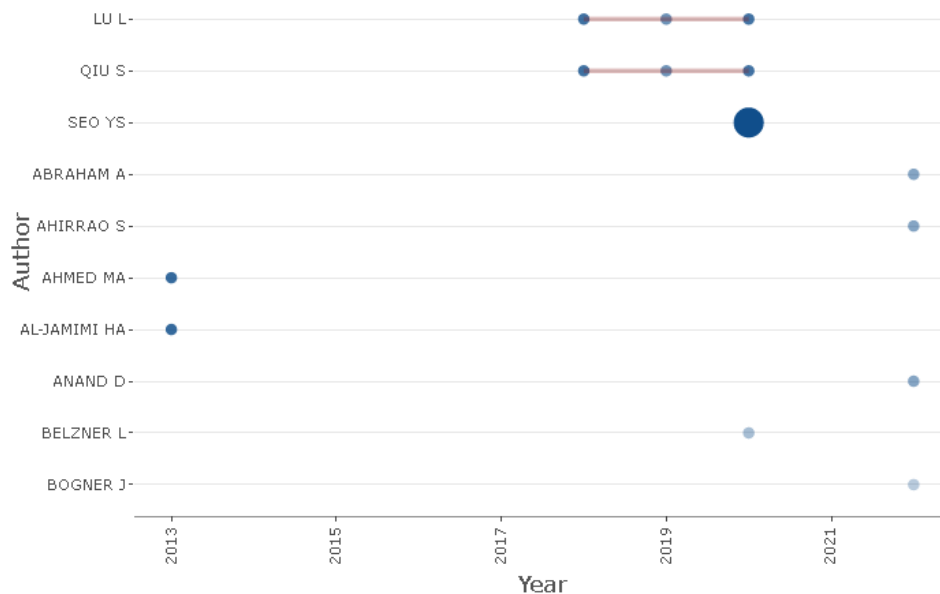
Figura 5 – países, autores e palavras-chave



Fonte: autoria própria

A Figura 6 ilustra a relação dos autores mais produtivos durante o período analisado. Graficamente, o tamanho do círculo é diretamente proporcional ao número de artigos publicados, enquanto a intensidade da cor do círculo representa o número de citações por ano. Assim, quanto maior o círculo, maior a quantidade de documentos publicados; quanto menor a transparência da cor do círculo, maior o número de citações do autor em determinado ano. A reta vermelha representa o período no qual os autores produziram material científico. Assim, pode-se concluir que o autor Lu, L., o mais produtivo, é um pesquisador recente e recebeu citações entre 2017 e 2020, mesmo período em que esteve em atividade.

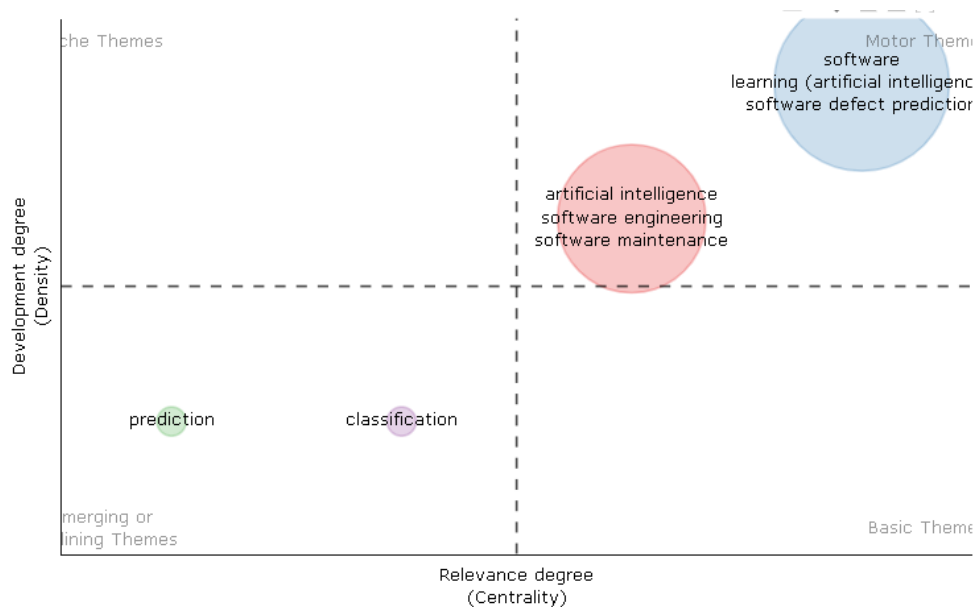
Figura 6 – autores e produção



Fonte: autoria própria

A análise do mapa temático (ou diagrama de co-palavras) distribui as palavras-chave em diferentes quadrantes de acordo com sua ocorrência. São considerados temas cuja densidade e centralidade são utilizadas em sua classificação e no mapeamento em um diagrama bidimensional. A Figura 7 apresenta o mapa temático da amostra. Os temas são divididos em 4 quadrantes: (1) quadrante superior direito: temas motores; (2) quadrante inferior direito: temas básicos; (3) quadrante inferior esquerdo: temas emergentes; (4) quadrante superior esquerdo: temas especializados ou de nicho. Neste caso, os temas de maior interesse são os motores e de nicho. Os temas motores são os mais explorados e que atualmente orientam as pesquisas. Os temas de nicho são especializados e detêm maior potencial para exploração, com baixas quantidades de bibliografia desenvolvida. Com base na Figura 7 é possível observar que os temas motores estão relacionados a inteligência artificial e software (engenharia, detecção de defeitos e manutenção). Não foram identificados temas básicos ou de nicho. Temas emergentes estão relacionados a predição e classificação.

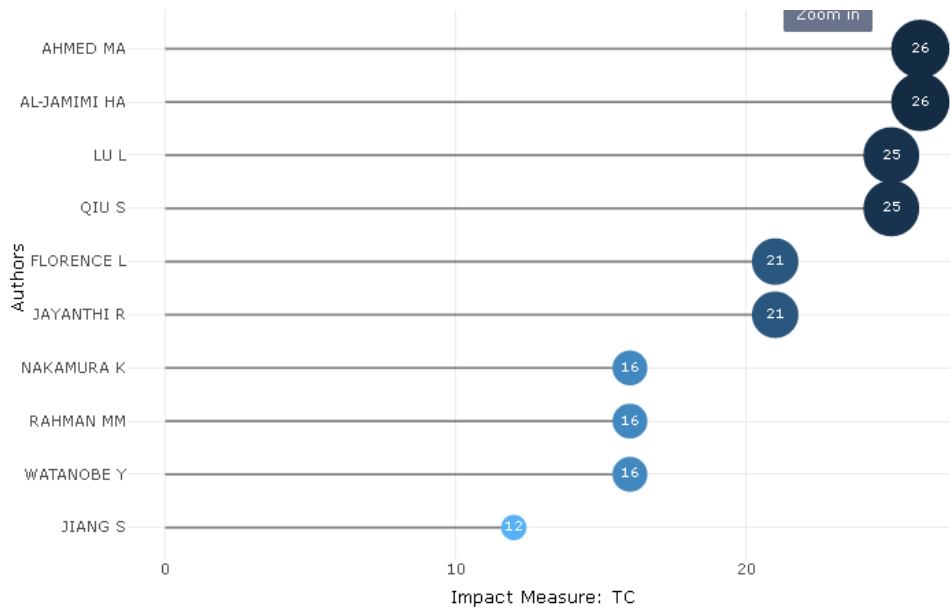
Figura 7 – mapa temático



Fonte: autoria própria.

A Figura 8 apresenta o impacto local dos autores com base no número de citações. Os mais influentes, com 72 citações cada um, são Ahmed, M. e Al-Jamimi, H.

Figura 8 – autores mais influentes por número de citações



Fonte: autoria própria

Pode-se inferir, referente à resposta à questão de pesquisa 1 (QP1), que foram publicados 19 artigos que estudam Inteligência Artificial aplicada à manutenção de software. Os artigos foram publicados a partir de 2013 e existe uma produção contemporânea em andamento, com três artigos já produzidos em 2022, até o mês de agosto. O principal *journal* que trata do assunto é o IET Software, com 8 publicações. Os autores mais influentes são Ahmed e Al-Jamimi. A maior comunidade de pesquisa está localizada na China. Os principais tópicos relacionados são classificados como motores; ou seja, estão no foco da pesquisa científica.

A partir da leitura dos artigos selecionados através do processo de amostragem, foi feita uma análise de conteúdo e codificação dos conceitos identificados. A Tabela 3 apresenta os resultados obtidos: *machine learning* é citado por todos os artigos e a manutenção preditiva é a mais citada.

Tabela 3 – análise de conteúdo

Código	Descrição	Ocorrências	Porcentagem
Inteligência artificial	<i>Machine learning</i>	Em todos os artigos	
	<i>Deep learning</i>	9	47%
	<i>Representation learning</i>	1	5%
	Total de artigos	19	
Manutenção	Corretiva	1	5%
	Evolutiva	2	11%
	Preditiva	17	19%
	Total de artigos	19	

Fonte: autoria própria

A questão de pesquisa 2 (QP2), conforme apresentado anteriormente no artigo, pode ser respondida através desta análise de conteúdo e codificação dos conceitos. Tal questão refere-se a quais são as aplicações de Inteligência Artificial utilizadas nas diferentes classificações de manutenção de software. Algumas observações podem ser feitas através da Tabela 3. Foram identificadas três aplicações de Inteligência Artificial: *machine learning*, *deep learning* e *representation learning*. *Machine learning* é a mais significativa, recebendo citações em todos os artigos. *Deep learning* está em segundo lugar, citado em 9 referências (47% da amostra). *Representation learning* é apresentado em apenas 1 artigo (5%) da amostra. As diferentes classificações de manutenção de software identificadas na análise dos artigos são corretiva, evolutiva e preditiva. A manutenção preditiva é a mais citada (17 artigos ou 89% da amostra). Em seguida, a manutenção evolutiva é mencionada em dois artigos (11% da amostra) e, finalmente, a manutenção corretiva figura em um artigo (5% da amostra).

5. DISCUSSÕES

Os resultados estão organizados segundo as categorias identificadas e analisadas. Em cada tópico é realizada a discussão dos resultados obtidos por meio da análise de

conteúdo. Uma tabela é apresentada ao término da seção a fim de ilustrar a distribuição das referências de acordo com os tópicos.

Os artigos selecionados foram inicialmente analisados com a finalidade de identificar a frequência de apresentação dos principais termos “inteligência artificial” e “manutenção”. As categorias referentes a Inteligência Artificial estudadas foram “*Artificial Intelligence*”, “*Machine Learning*”, “*Representation Learning*” e “*Deep Learning*” (BORGES et al., 2021). Todos os artigos citaram o termo “*machine learning*”; nove artigos citaram “*deep learning*” e apenas um artigo citou “*representative learning*”. Com relação à Manutenção, apenas um artigo referenciou “manutenção corretiva”, dois artigos citaram “manutenção evolutiva” e dezessete artigos tratam de “manutenção preditiva”. Este estudo indica que “*machine learning*” e “manutenção preditiva” são campos de pesquisa já explorados e que podem pertencer ao grupo de “temas motores”, conforme ilustrado anteriormente pela Figura 7. As Tabelas 4 e 5 sumarizam a distribuição das referências bibliográficas por tais classificações de manutenção e inteligência artificial.

Tabela 4 – análise de conteúdo: manutenção

Corretiva	Evolutiva	Preditiva
(LEE; SEO, 2020)	(GABOR et al., 2020), (HUH; HWA; SEO, 2020)	(JAYANTHI; FLORENCE, 2019), (PACHOULY et al., 2022), (PANDIT et al., 2022), (ZHANG; JING; WU, 2018), (QIU; LU; JIANG, 2018), (ZHU et al., 2020), (DENG; LU; QIU, 2020), (PANDEY; RATHEE; TRIPATHI, 2020), (KHAN, 2014), (CAI; LU; QIU, 2019), (WEI et al., 2019), (MARTÍNEZ-FERNÁNDEZ et al., 2022), (RAHMAN; WATANOBÉ; NAKAMURA, 2020), (HUH; HWA; SEO, 2020), (AHMED; AL-JAMIMI, 2013), (GALLI; CHICLANA; SIEWE, 2021), (WU; LIU; MA, 2018).

Fonte: autoria própria

Tabela 5 – inteligência artificial e componentes

Inteligência artificial	Machine learning	Representation learning	Deep learning
(GABOR et al., 2020), (JAYANTHI; FLORENCE, 2019), (PACHOULY et al., 2022), (PANDIT et al., 2022), (ZHANG; JING; WU, 2018), (QIU; LU; JIANG, 2018), (ZHU et al., 2020), (DENG; LU; QIU, 2020), (PANDEY; RATHEE; TRIPATHI, 2020), (KHAN, 2014), (CAI; LU; QIU, 2019), (WEI et al., 2019), (LEE; SEO, 2020), (MARTÍNEZ-FERNÁNDEZ et al., 2022), (RAHMAN; WATANOBE; NAKAMURA, 2020), (HUH; HWA; SEO, 2020), (AHMED; AL-JAMIMI, 2013), (GALLI; CHICLANA; SIEWE, 2021), (WU; LIU; MA, 2018)	(GABOR et al., 2020), (JAYANTHI; FLORENCE, 2019), (PACHOULY et al., 2022), (PANDIT et al., 2022), (ZHANG; JING; WU, 2018), (QIU; LU; JIANG, 2018), (ZHU et al., 2020), (DENG; LU; QIU, 2020), (PANDEY; RATHEE; TRIPATHI, 2020), (KHAN, 2014), (CAI; LU; QIU, 2019), (WEI et al., 2019), (LEE; SEO, 2020), (MARTÍNEZ-FERNÁNDEZ et al., 2022), (RAHMAN; WATANOBE; NAKAMURA, 2020), (HUH; HWA; SEO, 2020), (AHMED; AL-JAMIMI, 2013), (WU; LIU; MA, 2018), (GALLI; CHICLANA; SIEWE, 2021).	(ZHU et al., 2020)	(PACHOULY et al., 2022), (ZHU et al., 2020), (DENG; LU; QIU, 2020), (PANDEY; RATHEE; TRIPATHI, 2020), (CAI; LU; QIU, 2019), (WEI et al., 2019), (LEE; SEO, 2020), (MARTÍNEZ-FERNÁNDEZ et al., 2022), (RAHMAN; WATANOBE; NAKAMURA, 2020)

Fonte: autoria própria

A manutenção evolutiva é importante quando ocorrem mudanças no ambiente ou nas regras de negócio (MENS, 2016). Porém, a utilização da inteligência artificial como tecnologia de apoio é pouco explorada na literatura.

6. CONSIDERAÇÕES FINAIS

Este trabalho teve como principal objetivo analisar a literatura disponível referente às aplicações de inteligência artificial na manutenção de software. Utilizou-se a revisão de literatura para atingir tal objetivo e responder as questões de pesquisa indicadas com o uso de bibliometria e análise de conteúdo.

Os resultados obtidos a partir de métodos bibliométricos demonstram que o tema não é explorado ainda por um alto número de fontes. A primeira menção foi feita em 2013; foram publicados artigos em anos não consecutivos até o presente e o maior número de publicações (sete artigos) foi feito em 2020. Tais números podem representar as oportunidades a serem exploradas, já que os principais termos foram classificados como motores.

Quanto à resposta à primeira questão de pesquisa (estado da arte da literatura), pode-se afirmar que dois *journals* fizeram mais de uma publicação referente ao tema. A primeira é a IET Software, com oito publicações; a segunda é a *Applied Sciences – Basel*, com duas publicações. O autor mais produtivo, com três publicações, é Lu, L. Foi citado entre 2017 e 2020, mesmo período em que esteve em atividade. Quanto ao impacto, os autores com maior número de citações são Ahmed, A. e Al-Jamimi, A., ambos com vinte e seis citações. O país mais produtivo é a China, com cinco publicações.

A segunda questão de pesquisa (principais exemplos apresentados na literatura) é respondida através da verificação, através de uma análise de conteúdo, que a manutenção preditiva é a que mais emprega a inteligência artificial, especialmente através do uso de ferramentas de *machine learning*.



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

Com referência a limitações e direções de pesquisas futuras, identificou-se, durante a leitura das referências e construção deste artigo, que a disciplina de Qualidade de Software também utiliza amplamente a inteligência artificial como recurso, mas não foi avaliada neste trabalho. Identificou-se que a manutenção evolutiva, apesar de possuir importância, recebeu menores esforços de pesquisa até a atualidade. Desta forma, uma pesquisa aprofundada a respeito do tema é recomendada como caminhos futuros.

REFERÊNCIAS

14764-2021 - ISO/IEC/IEEE International Standard - Software engineering - Software life cycle processes - Maintenance. 2022.

15026-4-2021 - ISO/IEC/IEEE International Standard - Systems and software engineering--Systems and software assurance -- Part 4: Assurance in the life cycle. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9444258>>.

About Applied Sciences. Disponível em: <<https://www.mdpi.com/journal/applsci/about>>. Acesso em: 15 ago. 2022.

AHMED, M. A.; AL-JAMIMI, H. A. Machine learning approaches for predicting software maintainability: a fuzzy-based transparent model. **IET Software**, v. 7, n. 6, p. 317–326, dez. 2013.

BORGES, A. F. S. et al. The strategic use of artificial intelligence in the digital era: Systematic literature review and future research directions. **International Journal of Information Management**, v. 57, p. 102225, abr. 2021.

BOURQUE, P.; FAIRLEY, R. **Guide to the Software Engineering Body of Knowledge**. [s.l.] IEEE Computer Society, 2014.

CAI, Z.; LU, L.; QIU, S. An Abstract Syntax Tree Encoding Method for Cross-Project Defect Prediction. **IEEE Access**, v. 7, p. 170844–170853, 2019.

CARVALHO, M. M.; FLEURY, A.; LOPES, A. P. An overview of the literature on technology roadmapping (TRM): Contributions and trends. **Technological Forecasting and Social Change**, v. 80, n. 7, p. 1418–1437, set. 2013.

COBO, M. J. et al. Science mapping software tools: Review, analysis, and cooperative study among tools. **Journal of the American Society for Information Science and Technology**, v. 62, n. 7, p. 1382–1402, jul. 2011.



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

DENG, J.; LU, L.; QIU, S. Software defect prediction via LSTM. **IET Software**, v. 14, n. 4, p. 443–450, ago. 2020.

GABOR, T. et al. The scenario coevolution paradigm: adaptive quality assurance for adaptive systems. **International Journal on Software Tools for Technology Transfer**, v. 22, n. 4, p. 457–476, 6 ago. 2020.

GALLI, T.; CHICLANA, F.; SIEWE, F. Genetic Algorithm-Based Fuzzy Inference System for Describing Execution Tracing Quality. **Mathematics**, v. 9, n. 21, p. 2822, 6 nov. 2021.

HUH, J.-H.; HWA, J.; SEO, Y.-S. Hierarchical System Decomposition Using Genetic Algorithm for Future Sustainable Computing. **Sustainability**, v. 12, n. 6, p. 2177, 11 mar. 2020.

IEEE Standard Glossary of Software Engineering Terminology. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=159342>>.

IET Software Impact Factor & Key Scientometrics. Disponível em: <<https://www.scijournal.org/impact-factor-of-iet-softw.shtml>>. Acesso em: 15 ago. 2022.

JAYANTHI, R.; FLORENCE, L. Software defect prediction techniques using metrics based on neural network classifier. **Cluster Computing**, v. 22, n. S1, p. 77–88, 7 jan. 2019.

KHAN, M. J. Applications of case-based reasoning in Software Engineering: a systematic mapping study. **IET Software**, v. 8, n. 6, p. 258–268, dez. 2014.

LEE, D.-G.; SEO, Y.-S. Improving bug report triage performance using artificial intelligence based document generation model. **Human-centric Computing and Information Sciences**, v. 10, n. 1, p. 26, 16 dez. 2020.

MARTÍNEZ-FERNÁNDEZ, S. et al. Software Engineering for AI-Based Systems: A Survey. **ACM Transactions on Software Engineering and Methodology**, v. 31, n. 2, p. 1–59, 30 abr. 2022.

MENS, T. **An Ecosystemic and Socio-Technical View on Software Maintenance and Evolution**. 2016 IEEE International Conference on Software Maintenance and Evolution (ICSME). **Anais...IEEE**, out. 2016. Disponível em: <<http://ieeexplore.ieee.org/document/7816449/>>

PACHOULY, J. et al. A systematic literature review on software defect prediction using artificial intelligence: Datasets, Data Validation Methods, Approaches, and Tools. **Engineering Applications of Artificial Intelligence**, v. 111, p. 104773, maio 2022.



Anais do Simpósio Acadêmico de Engenharia de Produção (SAEPRO) da EEL-USP

VII SAEPRO – 14 e 15 de junho de 2023

PANDEY, S. K.; RATHEE, D.; TRIPATHI, A. K. Software defect prediction using K-PCA and various kernel-based extreme learning machine: an empirical study. **IET Software**, v. 14, n. 7, p. 768–782, 17 dez. 2020.

PANDIT, M. et al. Towards Design and Feasibility Analysis of DePaaS: AI Based Global Unified Software Defect Prediction Framework. **Applied Sciences**, v. 12, n. 1, p. 493, 4 jan. 2022.

QIU, S.; LU, L.; JIANG, S. Multiple-components weights model for cross-project software defect prediction. **IET Software**, v. 12, n. 4, p. 345–355, ago. 2018.

RAHMAN, MD. M.; WATANOBÉ, Y.; NAKAMURA, K. Source Code Assessment and Classification Based on Estimated Error Probability Using Attentive LSTM Language Model and Its Application in Programming Education. **Applied Sciences**, v. 10, n. 8, p. 2973, 24 abr. 2020.

RENTZ, S. M.; CARRINGTON, J. M.; BADGER, T. A. Two Strategies for Qualitative Content Analysis: An Intramethod Approach to Triangulation. **Qualitative Health Research**, v. 28, n. 5, p. 824–831, 9 abr. 2018.

SECINARO, S. et al. Employing bibliometric analysis to identify suitable business models for electric cars. **Journal of Cleaner Production**, v. 264, p. 121503, ago. 2020.

SOMMERVILLE, I. **Engenharia de Software**. São Paulo: Pearson Prentice Hall, 2011.

Systems and software engineering - Vocabulary. Disponível em:
<<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=5733835>>.

WEI, H. et al. Software Defect Prediction via Deep Belief Network. **Chinese Journal of Electronics**, v. 28, n. 5, p. 925–932, set. 2019.

WU, H.; LIU, H.; MA, Y. Empirical study on developer factors affecting tossing path length of bug reports. **IET Software**, v. 12, n. 3, p. 258–270, jun. 2018.

ZHANG, Z.-W.; JING, X.-Y.; WU, F. Low-rank representation for semi-supervised software defect prediction. **IET Software**, v. 12, n. 6, p. 527–535, dez. 2018.

ZHU, K. et al. Within-project and cross-project just-in-time defect prediction based on denoising autoencoder and convolutional neural network. **IET Software**, v. 14, n. 3, p. 185–195, jun. 2020.

ZUPIC, I.; ČATER, T. Bibliometric Methods in Management and Organization. **Organizational Research Methods**, v. 18, n. 3, p. 429–472, 22 jul. 2015.