

Mapeamento Sistemático sobre Métodos Computacionais para Análise das Trajetórias Profissionais de Egressos em Cursos de Computação e Áreas Correlatas

Camilla B. Quincozes, Alice Fonseca Finger, Aline Vieira de Mello

¹ Universidade Federal do Pampa (UNIPAMPA)
Curso de Engenharia de Software

{camillaborchhardt.aluno, alicefinger, alinemello}@unipampa.edu.br

Resumo. Este trabalho apresenta um mapeamento sistemático da literatura sobre estudos envolvendo egressos, com foco na identificação de fontes de dados, métodos analíticos, contribuições da Engenharia de Software, dimensões de trajetória e uso de técnicas de explicabilidade. Foram analisados cinquenta estudos primários. Os resultados indicam a predominância de questionários e estatística descritiva, além do crescimento do uso de dados de plataformas profissionais, como o LinkedIn. Contudo, observa-se baixa integração de fontes, uso limitado de técnicas de mineração e pouca atenção à explicabilidade, apontando lacunas e oportunidades para abordagens mais robustas e transparentes.

Abstract. This work presents a systematic mapping of the literature on studies involving graduates, focusing on the identification of data sources, analytical methods, contributions of Software Engineering, trajectory dimensions, and the use of explainability techniques. Fifty primary studies were analyzed. The results indicate the predominance of questionnaires and descriptive statistics, as well as the growing use of data from professional platforms, such as LinkedIn. However, low integration of sources, limited use of data mining techniques, and little attention to explainability are observed, pointing to gaps and opportunities for more robust and transparent approaches.

1. Introdução

O acompanhamento das trajetórias profissionais de egressos tem adquirido crescente relevância no contexto do ensino superior, especialmente em áreas tecnológicas como Computação. A expansão dos cursos e a demanda contínua por profissionais qualificados tornam essencial compreender como os graduados se inserem no mercado de trabalho, quais caminhos de carreira percorrem e quais competências desenvolvem ao longo da vida profissional. Estudos presentes na literatura reforçam essa necessidade ao analisarem empregabilidade, setores de atuação, progressão profissional e formação continuada [1, 2, 3, 4, 5].

Com a ampliação das fontes de dados nos últimos anos, investigações sobre egressos passaram a incorporar registros acadêmicos, questionários institucionais e, mais recentemente, plataformas profissionais como o *LinkedIn*. Diversos estudos [6, 7, 8, 9, 10] exploram perfis profissionais extraídos dessa rede para

identificar padrões de inserção, habilidades, mobilidade e redes profissionais, enquanto métodos tradicionais baseados em *surveys* permanecem amplamente utilizados. Em paralelo, começa a surgir o uso de técnicas de mineração de dados, processamento de texto e modelos supervisionados em contextos educacionais [11, 12, 13, 14, 15, 16, 17].

Apesar desses avanços, a literatura ainda se mostra fragmentada e heterogênea. Muitos estudos analisam apenas uma fonte de dados (por exemplo, *surveys* ou perfis em redes profissionais), limitando o potencial explicativo das análises. Além disso, observa-se baixa disponibilidade de artefatos reutilizáveis de Engenharia de Software, como pipelines, ferramentas ou *frameworks open-source*, o que dificulta a replicação e a comparação entre instituições. A adoção de técnicas de explicabilidade também é incipiente: a maior parte dos trabalhos opera com abordagens de *caixa-preta* e não discute de forma sistemática a transparência dos modelos utilizados [17, 16].

Nesse cenário, torna-se relevante organizar de forma sistemática o conhecimento produzido sobre trajetórias profissionais de egressos da área da Computação e de áreas correlatas. Este mapeamento sistemático tem como objetivo identificar, organizar e categorizar estudos que investigam essas trajetórias, considerando cinco dimensões centrais: (i) fontes de dados utilizadas; (ii) métodos estatísticos e computacionais empregados; (iii) contribuições da Engenharia de Software (ferramentas, processos e reusabilidade); (iv) dimensões de carreira analisadas (empregabilidade, progressão, impacto, competências); e (v) uso, ou ausência, de XAI.

Ao seguir um protocolo rigoroso de mapeamento sistemático, este estudo oferece uma visão estruturada e atualizada da literatura sobre egressos em Computação e áreas correlatas, identificando lacunas e oportunidades de pesquisa, especialmente quanto à integração de fontes de dados, ao uso de métodos computacionais mais avançados e à adoção de modelos explicáveis. Além disso, com base no protocolo de busca e seleção empregado, não foram identificados mapeamentos ou revisões secundárias com recorte equivalente na Educação em Computação, combinando métodos analíticos, fontes de dados, práticas de Engenharia de Software e explicabilidade. Assim, este trabalho se posiciona como uma contribuição estruturante para a área, ao consolidar evidências, organizar o domínio e apontar direções para pesquisas futuras.

Este artigo está organizado da seguinte forma. A Seção 2 descreve o protocolo para o mapeamento sistemático. A Seção 3 apresenta a condução do mapeamento. Os resultados são apresentados e discutidos na Seção 4. A Seção 5 discute as ameaças à validade do estudo. Por fim, a Seção 6 apresenta as conclusões.

2. Processo de Mapeamento Sistemático

O presente estudo segue as diretrizes para Mapeamentos Sistemáticos em ES propostas por Kitchenham e Petersen [18, 19]. O protocolo foi elaborado previamente e conduzido de forma estruturada e reproduzível. Esta seção apresenta o escopo, as questões de pesquisa, o processo de busca, os critérios de elegibilidade, o processo de seleção e os procedimentos de extração e análise dos dados.

Este mapeamento tem como objetivo identificar, organizar e categorizar estudos que investigam trajetórias profissionais de egressos da área de Computação e de áreas correlatas. O escopo foi definido para abarcar cinco dimensões: (i) fontes de dados utilizadas; (ii) métodos estatísticos e técnicas de mineração de dados; (iii) contribuições da Engenharia de Software; (iv) dimensões da trajetória profissional analisadas; e (v) adoção de técnicas de XAI.

As seguintes questões de pesquisa foram elaboradas para orientar a análise sistemática: **QP1:** Quais fontes de dados são utilizadas para investigar trajetórias de egressos em Computação e áreas correlatas? **QP2:** Quais técnicas estatísticas, computacionais ou de mineração de dados são empregadas? **QP3:** Como a Engenharia de Software contribui para soluções voltadas ao monitoramento de egressos? **QP4:** Quais dimensões da trajetória profissional são analisadas? **QP5:** Quais técnicas de explicabilidade são utilizadas? **QP6:** Quais lacunas e oportunidades futuras são identificadas?

A string de busca foi construída com base nas questões de pesquisa e adaptada conforme as restrições de cada repositório.

```
("alumni" OR "graduates" OR "egressos" OR "career trajectory" OR "employability") AND ("computer science" OR "computação") AND ("data mining" OR "skills" OR "linkedin" OR "rais" OR "job market")
```

A string foi refinada iterativamente com base em buscas exploratórias e nos termos recorrentes dos estudos recuperados, a fim de melhorar a aderência ao recorte da pesquisa. As buscas foram realizadas entre outubro e novembro de 2025 nas bases IEEE Xplore, ACM Digital Library, Google Scholar e arXiv, sendo esta última incluída para ampliar a cobertura de estudos recentes, incluindo *pre-prints* e trabalhos em fase de publicação. Os critérios de seleção foram organizados em duas categorias e são apresentados nas Tabelas 1 e 2.

Tabela 1. Critérios de Inclusão

ID	Descrição
I1	Estudos publicados entre 2000 e 2025.
I2	Escritos em inglês ou português.
I3	Estudos empíricos focados em trajetórias de egressos da área de Computação ou áreas tecnológicas correlatas.
I4	Uso de fontes como LinkedIn, RAIS, CNPJ, dados institucionais ou questionários.
I5	Aplicação de métodos estatísticos, computacionais, XAI ou contribuições de Engenharia de Software.
I6	Análise de dimensões como empregabilidade, progressão, liderança ou empreendedorismo.

Tabela 2. Critérios de Exclusão

ID	Descrição
E1	Estudos duplicados.
E2	Trabalhos com menos de cinco páginas efetivas.

O processo de seleção ocorreu em três etapas: (i) triagem por título e resumo para descarte de irrelevantes; (ii) leitura das metodologias para aplicação dos critérios de inclusão; e (iii) consolidação do conjunto final, com registro sistemático das decisões. O resultado foi um conjunto de estudos que atende simultaneamente ao escopo definido, aos critérios de elegibilidade e aos requisitos mínimos de qualidade estabelecidos neste protocolo. A seleção dos estudos

foi conduzida por três pesquisadores, com triagem por título e resumo, seguida de leitura em texto completo dos trabalhos potencialmente relevantes. As decisões de inclusão e exclusão foram tomadas por consenso, com base nos critérios definidos no protocolo.

Após a seleção final, cada estudo foi avaliado segundo cinco Questões de Qualidade (QQ): **QQ1**: Descrição clara da fonte de dados e público-alvo; **QQ2**: Transparência no método de análise; **QQ3**: Contribuição explícita de Engenharia de Software; **QQ4**: Uso de técnicas de explicabilidade; **QQ5**: Identificação e discussão das dimensões da trajetória profissional.

Cada estudo foi avaliado nas cinco questões de qualidade (QQ1–QQ5), com pesos distintos: QQ1 e QQ5 variam de 0 a 2,0 pontos; QQ2 varia de 0 a 1,5 ponto; e QQ3 e QQ4 variam de 0 a 1,0 ponto cada, totalizando um máximo de 7,5 pontos por estudo. A Tabela 3 apresenta os resultados completos da avaliação de qualidade. A coluna **Sc** corresponde à pontuação total obtida, enquanto a coluna **D** sintetiza essa pontuação em faixas qualitativas: **E** (Excelente, ≥ 6.5), **MB** (Muito Bom, 5.0–6.25), **B** (Bom, 3.5–4.75), **M** (Médio, 2.0–3.25) e **R** (Ruim, < 2.0). Essa classificação é utilizada posteriormente para interpretar a robustez metodológica do conjunto de estudos na discussão dos resultados.

Tabela 3. Notas da Avaliação de Qualidade (Pontuação Máxima: 7.5)

ID	Ref.	Ano	Q1	Q2	Q3	Q4	Q5	Sc	D	ID	Ref.	Ano	Q1	Q2	Q3	Q4	Q5	Sc	D
E01	[1]	22	2.0	1.5	1.0	0	2.0	6.5	E	E26	[20]	20	2.0	1.5	1.0	0	2.0	6.5	E
E02	[21]	19	2.0	1.5	0.5	0	0.5	4.5	B	E27	[5]	15	2.0	1.5	1.0	0	2.0	6.5	E
E03	[2]	10	2.0	0.8	0	0	0.5	3.3	M	E28	[9]	22	2.0	1.5	0.5	0.5	2.0	6.5	E
E04	[8]	23	2.0	1.5	1.0	0	0.5	5.0	MB	E29	[22]	21	2.0	1.5	0.5	0.5	2.0	6.5	E
E05	[7]	23	2.0	1.5	1.0	0.5	0.5	5.5	MB	E30	[15]	22	2.0	1.5	0.5	0.5	2.0	6.5	E
E06	[3]	11	2.0	0.8	0	0	0	2.8	M	E31	[23]	21	2.0	1.5	0.5	0	2.0	6.0	MB
E07	[24]	20	2.0	0.8	0.5	0	2.0	5.3	MB	E32	[25]	20	2.0	1.5	0.5	0.5	2.0	6.5	E
E08	[26]	16	2.0	1.5	1.0	0	0	4.5	B	E33	[27]	23	2.0	1.5	1.0	1.0	2.0	7.5	E
E09	[28]	21	2.0	1.5	1.0	1.0	2.0	7.5	E	E34	[29]	19	2.0	1.5	1.0	0.5	2.0	7.0	E
E10	[6]	24	2.0	1.5	0.5	0	0	4.0	B	E35	[30]	23	2.0	1.5	1.0	0.5	2.0	7.0	E
E11	[13]	16	2.0	1.5	1.0	0	2.0	6.5	E	E36	[14]	25	2.0	1.5	1.0	0.5	2.0	7.0	E
E12	[31]	20	2.0	1.5	1.0	0.5	2.0	7.0	E	E37	[32]	25	2.0	1.5	1.0	0.5	2.0	7.0	E
E13	[33]	15	0.5	0.5	0	0	0.5	1.5	R	E38	[34]	15	2.0	0.8	1.5	0	2.0	6.3	MB
E14	[12]	22	1.0	0.8	0	0	0	1.8	R	E39	[35]	24	2.0	1.5	0.5	1.0	2.0	7.0	E
E15	[4]	23	1.0	0	0.5	0	0	1.5	R	E40	[36]	25	2.0	1.5	1.0	0.5	2.0	7.0	E
E16	[37]	23	2.0	1.5	0.5	0	1.0	5.0	MB	E41	[38]	20	2.0	0.8	1.0	0	2.0	5.8	MB
E17	[39]	13	2.0	0.8	0	0	0.5	3.3	M	E42	[40]	17	2.0	1.5	1.0	0.5	2.0	7.0	E
E18	[17]	24	2.0	1.5	0	0	1.0	4.5	B	E43	[41]	16	2.0	1.5	0.8	0.5	2.0	6.8	E
E19	[42]	15	2.0	1.5	1.0	0	0.5	5.0	MB	E44	[43]	24	2.0	1.5	1.0	0.5	2.0	7.0	E
E20	[44]	23	2.0	1.5	1.0	0	1.0	5.5	MB	E45	[45]	22	2.0	1.5	1.0	0.5	2.0	7.0	E
E21	[10]	19	2.0	1.0	0.5	0	0.5	4.0	B	E46	[46]	20	2.0	1.5	1.0	0	0	4.5	B
E22	[11]	23	2.0	0.8	0.5	0	0.5	3.8	B	E47	[47]	25	2.0	1.5	1.0	0.5	2.0	7.0	E
E23	[48]	24	2.0	0.8	0.5	0	0.5	3.8	B	E48	[49]	16	2.0	1.5	0	0	0	3.5	B
E24	[16]	25	2.0	1.5	0	0	0	3.5	B	E49	[50]	22	1.0	0	1.0	0	0	2.0	M
E25	[51]	25	2.0	1.5	1.0	0	1.0	5.5	MB	E50	[52]	23	2.0	1.5	0.5	1.0	2.0	7.0	E

Legenda: Q1/Q5: Max 2.0; Q2: Max 1.5; Q3/Q4: Max 1.0. **Sc:** Score Total. **D (Descrição):** E: Excelente (≥ 6.5), MB: Muito Bom (5.0 – 6.25), B: Bom (3.5 – 4.75), M: Médio (2.0 – 3.25), R: Ruim (< 2.0).

*Nomes e anos abreviados para ajuste de layout. 0.75 arredondado visualmente para 0.8.

3. Condução

As buscas realizadas nas quatro bases definidas no protocolo resultaram em 502 estudos identificados. A distribuição por fonte foi:

- IEEE Xplore: 6 estudos;
- ACM Digital Library: 464 estudos;
- Google Scholar: 10 estudos;
- arXiv: 22 estudos.

Após a remoção de 19 estudos duplicados, 483 artigos foram submetidos à triagem por título e resumo, resultando em 101 estudos pré-selecionados. A leitura direcionada das seções de metodologia levou à seleção final de 50 estudos, enquanto 51 foram excluídos por não atenderem aos critérios de elegibilidade estabelecidos no protocolo. A Tabela 4 apresenta o desempenho agregado dos 50 estudos selecionados, refletindo o grau de clareza metodológica, contribuição técnica e maturidade analítica observados.

Tabela 4. Resultados agregados da avaliação de qualidade dos estudos selecionados

Questão de Qualidade	Nota Máxima	Nota Média	Nota Zero
QQ1 – Fonte de dados e público-alvo bem descritos	46	4	0
QQ2 – Método de análise apresentado de forma transparente	37	11	2
QQ3 – Contribuição explícita de Engenharia de Software	26	16	8
QQ4 – Uso de técnicas de explicabilidade	4	16	30
QQ5 – Dimensões da trajetória identificadas e discutidas	27	14	9

4. Resultados e Discussão

Nesta seção, são apresentados e discutidos os resultados da análise dos estudos primários, organizados de acordo com cada questão de pesquisa.

4.1. QP1: Quais fontes de dados são utilizadas para investigar trajetórias de egressos em Computação e áreas correlatas?

A análise dos estudos primários revelou que as fontes de dados empregadas na investigação de trajetórias de egressos podem ser organizadas em cinco categorias, diferenciadas principalmente pela origem e pela estratégia de coleta: (i) coleta ativa via questionários/surveys, (ii) extração de perfis e interações em plataformas profissionais, (iii) bases institucionais, (iv) integrações híbridas e (v) bases governamentais e portais de emprego. A Tabela 5 apresenta a distribuição detalhada das fontes utilizadas. De forma geral, essa distribuição quantitativa evidencia a manutenção de abordagens tradicionais, ao mesmo tempo em que aponta um avanço gradual de estratégias automatizadas de coleta na web, motivadas por escala e atualização mais frequente dos registros.

Diferentemente da hipótese inicial de predominância absoluta de redes sociais, questionários e *surveys* constituem a fonte mais recorrente (17 estudos), sendo especialmente adequados para capturar dimensões subjetivas da trajetória profissional, como satisfação, percepção de competências e expectativas de carreira. No entanto, esse tipo de coleta tende a sofrer com limitações conhecidas,

Tabela 5. Fontes de dados utilizadas nos estudos selecionados

Fonte de dados	Qtd. Estudos	IDs dos Artigos
Questionários / Surveys	17	E03, E05, E07, E09, E13, E14, E17, E22, E23, E28, E31, E32, E33, E37, E40, E48, E50
LinkedIn (Dados de perfis públicos)	8	E10, E11, E12, E19, E21, E27, E41, E43
Dados Institucionais (Sistemas Acadêmicos)	8	E02, E04, E06, E16, E34, E38, E44, E47
Fontes Híbridas (Integração de fontes)	7	E08, E20, E24, E30, E35, E39, E45
Bases Gov. (RAIS/CNPJ) e Portais de Vagas	6	E01, E25, E26, E29, E36, E42
Revisões/Propostas	4	E15, E18, E46, E49

como viés de seleção, baixa taxa de resposta e defasagem temporal, o que reforça a importância de complementar a análise com fontes observacionais.

Nesse sentido, o *LinkedIn* consolidou-se como a principal fonte digital, aparecendo como fonte exclusiva em oito estudos e compondo parte expressiva das abordagens híbridas. Seu uso está fortemente associado a técnicas de *web scraping* e à construção de *datasets* em larga escala, viabilizando análises sobre cargos, progressão e *skills*. Ainda assim, desafios de acesso, termos de uso e variações na estrutura dos dados impõem barreiras de reprodutibilidade, frequentemente relatadas de forma indireta nos estudos.

Bases governamentais e portais de emprego foram fontes menos exploradas, geralmente circunscritas a contextos regionais, seja pela dificuldade técnica de acesso aos dados, seja por impedimentos legais. Embora essas fontes proporcionem uma visão privilegiada sobre a formalização dos vínculos e a distribuição por setores econômicos, elas tendem a ser menos ricas na descrição detalhada das competências e das atividades reais exercidas pelos profissionais.

Por fim, destaca-se a categoria de fontes híbridas, que integra, por exemplo, registros acadêmicos institucionais com dados externos do *LinkedIn* ou resultados de questionários. Embora mais complexas do ponto de vista de Engenharia de Software, por demandarem integração de dados, limpeza, normalização e resolução de entidades, essas abordagens mostram maior potencial explicativo ao correlacionar desempenho acadêmico com desfechos profissionais subsequentes. A baixa incidência de estudos que exploram bases governamentais em conjunto com dados institucionais indica uma oportunidade relevante para pesquisas futuras voltadas à integração de dados abertos, governança e reprodutibilidade.

4.2. QP2: Quais técnicas de mineração de dados e métodos analíticos são aplicados?

A Tabela 6 sintetiza a distribuição das técnicas identificadas, permitindo uma visão geral das abordagens adotadas nos estudos analisados. Essa distribuição revela um panorama onde abordagens tradicionais coexistem com métodos computacionais avançados, evidenciando diferentes níveis de maturidade analítica no campo. Predominam, contudo, a estatística descritiva e a visualização de dados, especialmente em investigações fundamentadas em questionários e estudos de caso, cujo objetivo central reside na caracterização de perfis de egressos e na

descrição de padrões de inserção profissional.

Tabela 6. Técnicas de análise e mineração de dados utilizadas nos estudos

Técnica / Abordagem	Qtd. Estudos	IDs dos Artigos
Estatística Descritiva / Visualização	21	E03, E06, E07, E11, E13, E14, E16, E17, E21, E24, E26, E29, E30, E31, E32, E34, E41, E43, E45, E48, E50
ML Supervisionado (Classificação/Predição)	11	E04, E05, E09, E33, E35, E37, E39, E40, E42, E44, E47
Mineração de Dados Não-Superv. (Cluster/Assoc.)	6	E02, E12, E19, E20, E27, E28
Mineração de Texto (NLP) / Topic Modeling	4	E01, E08, E25, E36
Análise Qualitativa / Temática	3	E22, E23, E38

Simultaneamente, um subconjunto expressivo de pesquisas adota técnicas de aprendizado de máquina supervisionado, com foco em tarefas de predição de empregabilidade, recomendação de carreiras e identificação de perfis. A recorrência de algoritmos como *Árvores de Decisão*, *Random Forest* e *Support Vector Machine* (SVM) sugere uma preferência por modelos consolidados e, frequentemente, interpretáveis. Tais abordagens permitem explorar relações complexas entre variáveis acadêmicas, profissionais e demográficas, ampliando significativamente o potencial explicativo das análises.

De maneira mais pontual, emergem a mineração de texto e as técnicas de Processamento de Linguagem Natural (NLP), geralmente aplicadas à extração automática de competências e à análise de descrições de vagas ou currículos. Embora menos frequentes, esses métodos constituem um eixo promissor para trabalhos futuros, pois possibilitam alinhar currículos acadêmicos às demandas reais do mercado. Complementarmente, abordagens não supervisionadas, como *clustering* e regras de associação, e análises qualitativas aparecem como suporte, sobretudo em estudos de caráter exploratório.

Os dados confirmam que, embora os modelos preditivos representem a principal fronteira de inovação técnica, sua adoção ainda ocorre em escala reduzida quando comparada aos métodos descritivos. Nota-se, ainda, que a aplicação de técnicas de explicabilidade associadas a esses modelos permanece incipiente, o que pode limitar a compreensão dos fatores determinantes nas predições e, conseqüentemente, a confiança dos usuários nas soluções propostas.

4.3. QP3: Como a Engenharia de Software contribui para soluções de monitoramento de egressos?

A análise dos estudos primários indica que a Engenharia de Software desempenha um papel transversal no monitoramento de egressos, contribuindo tanto para a fundamentação empírica do domínio quanto para o desenvolvimento de soluções técnicas. Essas contribuições podem ser organizadas em quatro categorias principais, sintetizadas na Tabela 7, que refletem diferentes níveis de maturidade e orientação prática dos trabalhos.

Uma parcela significativa dos estudos concentra-se na geração de diagnósticos e *insights* empíricos, os quais, embora não resultem diretamente em artefatos

Tabela 7. Classificação das contribuições de Engenharia de Software

Tipo de Contribuição	Qtd	IDs dos Artigos
Insights Empíricos / Requisitos (Diagnósticos)	20	E03, E06, E07, E13, E14, E16, E17, E22, E23, E24, E29, E30, E31, E32, E38, E43, E45, E46, E48, E50
Modelos Preditivos / Algoritmos (ML)	10	E04, E05, E12, E25, E33, E35, E39, E40, E42, E44
Sistemas, Ferramentas e Protótipos	10	E01, E09, E15, E26, E34, E36, E37, E41, E47, E49
Metodologias, Pipelines e Arquiteturas	10	E02, E08, E10, E11, E18, E19, E20, E21, E27, E28

de software, são importantes por subsidiarem a definição de requisitos de domínio. Esses trabalhos identificam variáveis relevantes como características acadêmicas, demográficas e profissionais que orientam o projeto de sistemas de acompanhamento e apoio à tomada de decisão institucional.

No que se refere às contribuições técnicas, observa-se um equilíbrio entre o desenvolvimento de modelos preditivos baseados em aprendizado de máquina e a implementação de sistemas e ferramentas aplicadas, como portais web e *dashboards* de monitoramento. Enquanto os modelos enfatizam a acurácia e a capacidade preditiva, os sistemas tendem a priorizar a usabilidade e a aplicação prática em contextos institucionais, aproximando os resultados analíticos de gestores e formuladores de políticas educacionais.

Além disso, alguns estudos propõem metodologias, pipelines e arquiteturas para coleta, integração e análise de dados de egressos, frequentemente envolvendo a extração automatizada de informações de plataformas profissionais. Essas contribuições também refletem aspectos de Engenharia de Software, especialmente no que se refere à modularidade, ao reuso e à organização do fluxo de dados. Contudo, a avaliação de qualidade revela limitações recorrentes quanto à documentação técnica e à disponibilização pública dos artefatos desenvolvidos, o que compromete a replicabilidade e o reuso das soluções propostas.

De forma geral, os resultados sugerem que, embora a Engenharia de Software, esteja presente de maneira consistente no campo, ainda predomina a adoção de soluções *ad-hoc* e pouco padronizadas. A ausência de arquiteturas de referência e de ferramentas consolidadas indica uma oportunidade de pesquisa voltada à sistematização de práticas, à abertura de artefatos e ao fortalecimento da reprodutibilidade em sistemas de monitoramento de trajetórias de egressos.

4.4. QP4: Quais dimensões da trajetória profissional são analisadas?

A extração e categorização dos dados permitiram identificar cinco dimensões analíticas recorrentes nos estudos sobre trajetórias profissionais dos egressos: empregabilidade, competências, progressão de carreira, aspectos sociais e mobilidade ou empreendedorismo (Tabela 8). De forma geral, observa-se que a literatura se afasta gradualmente de uma visão estritamente centrada no status de emprego, passando a incorporar dimensões relacionadas às competências e ao alinhamento entre formação acadêmica e demandas do mercado.

Tabela 8. Dimensões da trajetória profissional analisadas nos estudos

Dimensão Analisada	Qtd. Estudos	IDs dos Artigos
Empregabilidade (Status, Área, Vínculo)	18	E02, E03, E04, E06, E13, E14, E15, E17, E21, E24, E26, E32, E38, E39, E41, E46, E49, E18
Competências e Habilidades (Skills)	13	E01, E05, E08, E20, E22, E34, E36, E37, E40, E42, E44, E48, E50
Progressão e Evolução (Tempo/Cargos)	10	E09, E11, E12, E19, E23, E27, E28, E35, E43, E45
Aspectos Sociais (Gênero/Demografia)	5	E07, E16, E25, E30, E33
Mobilidade e Empreendedorismo	4	E10, E29, E31, E47

Predominam os estudos focados em empregabilidade e competências, indicando uma ênfase em indicadores mais imediatos e mensuráveis, frequentemente utilizados para fins de prestação de contas institucional e avaliação de cursos. A análise de competências, em particular, tem sido impulsionada pelo uso de técnicas de mineração de texto, que permitem identificar automaticamente o vão entre os conteúdos curriculares e as exigências do mercado de trabalho.

Em contraste, dimensões de caráter longitudinal, como progressão e evolução de carreira, aparecem com menor frequência, embora sejam essenciais para a compreensão de trajetórias profissionais ao longo do tempo. Esses estudos tendem a analisar mudanças de cargos, tempo para promoção e permanência na área, oferecendo uma perspectiva mais dinâmica, porém ainda pouco explorada no conjunto da literatura.

Aspectos sociais e demográficos, como gênero e fatores socioeconômicos, bem como mobilidade geográfica e empreendedorismo, constituem as dimensões menos investigadas. A baixa incidência desses recortes sugere limitações na compreensão dos impactos sociais e regionais da formação em Computação e áreas Correlatas, além de indicar oportunidades para pesquisas futuras que articulem trajetórias profissionais a contextos territoriais, diversidade e desenvolvimento econômico.

QP5: Quais técnicas de explicabilidade são utilizadas?

Conforme apresentado na Tabela 9, a distribuição das abordagens identificadas evidencia que a explicabilidade permanece como uma das principais lacunas na literatura sobre o monitoramento e a análise de trajetórias de egressos.

No conjunto analisado, uma parcela expressiva dos estudos não apresenta mecanismos explícitos para explicar resultados, relações identificadas ou previsões produzidas por modelos, operando sob uma lógica de *caixa-preta*. Esse cenário é particularmente sensível em aplicações que suportam recomendações de carreira e tomada de decisão institucional, pois a falta de transparência pode reduzir a confiança, dificultar a auditoria e limitar a adoção das soluções propostas.

Quando a explicabilidade aparece, ela ocorre predominantemente de forma indireta, por meio da escolha de modelos intrinsecamente interpretáveis. Estu-

Tabela 9. Estratégias de explicabilidade (XAI) identificadas nos estudos

Nível de Explicabilidade	Descrição da Abordagem	IDs dos Estudos
Ausência (Black-box)	Foco exclusivo em métricas de desempenho ou descrição de dados, sem mecanismos de explicação causal.	E01, E03, E04, E06, E07, E08, E10, E11, E13, E14, E15, E16, E17, E18, E24, E26, E29, E30, E31, E32, E34, E36, E37, E40, E41, E42, E43, E44, E45, E46, E47, E48, E49
Intrínseca (White-box)	Uso deliberado de modelos interpretáveis (Árvores, Regras, Regressão, Fatores).	E02, E05, E12, E19, E20, E21, E25, E27, E28, E33, E35, E39, E50
Pós-Hoc (Model-Agnostic)	Aplicação de técnicas externas (LIME) para explicar predições.	E09

dos que utilizam Árvores de Decisão, Regressão Logística, Regras de Associação e Métodos de Clustering permitem que a lógica decisória (ou a formação dos grupos) seja inspecionada diretamente a partir da estrutura do modelo. Entretanto, mesmo nesses casos, observa-se que a discussão frequentemente se limita a interpretações pontuais, sem uma estratégia sistemática de explicação orientada a usuários finais.

O uso de técnicas *pós-hoc*, independentes de modelo, é raro no conjunto analisado, aparecendo de forma isolada com o uso do *LIME* para justificar predições de modelos mais complexos. A baixa adoção de abordagens consolidadas de XAI reforça uma oportunidade concreta de pesquisa: desenvolver pipelines reproduzíveis que combinem modelos preditivos com mecanismos de explicação compreensíveis.

4.5. QP6: Quais lacunas e oportunidades futuras são identificadas?

A síntese dos estudos, combinada com a avaliação de qualidade, permitiu identificar cinco lacunas estruturais que delimitam oportunidades de pesquisa e desenvolvimento na área. Em primeiro lugar, observa-se a predominância de abordagens de *caixa-preta*, com baixa incorporação de mecanismos de explicabilidade, o que compromete transparência, auditoria e confiança em aplicações sensíveis, como recomendação de carreira e apoio à tomada de decisão institucional. Em segundo lugar, há uma concentração em análises estáticas centradas em empregabilidade, enquanto investigações de natureza longitudinal que acompanhem progressão, transições e estabilidade permanecem minoritárias.

A terceira lacuna refere-se aos silos de dados: a maioria dos estudos utiliza uma única fonte, seja a percepção do egresso via *survey*, seja dados observacionais extraídos de plataformas digitais. A integração híbrida de fontes, embora mais custosa em termos de Engenharia de Software, ainda é pouco frequente, apesar de seu maior potencial explicativo. Em quarto lugar, destaca-se a volatilidade e baixa reutilização das ferramentas propostas: mesmo quando artefatos são desenvolvidos, frequentemente faltam documentação, disponibilização pública e padronização, limitando replicabilidade e reuso. Por fim, identifica-se ausência de consenso sobre métricas e definições de sucesso profissional, com estudos utilizando critérios distintos (por exemplo, salário, aderência à área de formação, estabilidade ou cargos), dificultando comparação direta entre resultados e a consolidação de indicadores compartilhados.

5. Ameaças à Validade

A validade de conclusão refere-se à precisão dos resultados sintetizados e à confiabilidade da extração dos dados. Uma ameaça identificada durante a análise foi a baixa transparência metodológica de parte dos estudos primários. Vários trabalhos, especialmente aqueles focados apenas em resultados de ferramentas, não detalharam os hiperparâmetros dos algoritmos ou as métricas de validação, dificultando a classificação precisa das técnicas na QP2. Para mitigar esse risco, a extração priorizou informações explícitas e, nos casos de ambiguidade, classificou-se a técnica em categorias mais amplas, por exemplo *Machine Learning* supervisionado em vez de um algoritmo específico, evitando inferências incorretas.

A validade interna diz respeito às influências que podem ter afetado o processo de seleção e análise. Uma ameaça relevante foi a heterogeneidade das fontes de dados. A comparação entre estudos que utilizam bases massivas e estruturadas e estudos baseados em questionários subjetivos com amostras pequenas, pode introduzir viés na síntese dos resultados. Para mitigar essa ameaça, a análise dos dados foi segmentada: os resultados quantitativos foram tratados separadamente das dimensões qualitativas, garantindo que conclusões sobre *Big Data* não fossem contaminadas por limitações de amostra de estudos menores.

A validade de construto está relacionada à adequação das definições utilizadas ao fenômeno investigado. A principal ameaça detectada nesta revisão foi a polissemia do termo trajetória. A análise da QP4 revelou que, embora muitos estudos utilizem o termo em seus títulos, na prática, cerca de 36% analisaram apenas a Empregabilidade, sem observar a evolução longitudinal, como promoções e mobilidade, que caracterizam uma trajetória. Para reduzir o impacto dessa discrepância conceitual, adotou-se um esquema de classificação rigoroso na extração, diferenciando explicitamente estudos estáticos de estudos dinâmicos, garantindo que as lacunas apontadas na QP6 refletissem a ausência real de análises longitudinais e não apenas uma questão terminológica.

A validade externa refere-se ao grau de generalização dos resultados. Uma limitação inerente aos estudos primários selecionados é o foco regional e institucional. Grande parte dos trabalhos, especialmente aqueles classificados como Estudos de Caso na extração, limitou-se a analisar egressos de uma única universidade ou de um contexto geográfico específico, com forte presença de estudos do Brasil, China e EUA. Portanto, os padrões de carreira e as competências mais demandadas identificados nesta revisão podem não ser universalmente generalizáveis para todos os contextos. A predominância de relatos de experiência locais sugere que os resultados representam um panorama das práticas de monitoramento existentes, mas não necessariamente uma verdade universal sobre as trajetórias em si. As estratégias adotadas contribuíram para minimizar os riscos. No entanto, ainda é fundamental interpretar os resultados considerando que a literatura atual ainda é fragmentada e fortemente dependente de contextos locais.

O conjunto de dados completo contendo a extração e classificação de todos os estudos analisados neste trabalho foi depositado publicamente no repositório *Zenodo* e pode ser acessado através deste link permanente.

6. Conclusões

Este mapeamento sistemático analisou, de forma estruturada, como a literatura nacional e internacional tem investigado as trajetórias profissionais de egressos da área da Computação e áreas correlatas. A síntese dos 50 estudos indica um campo em consolidação, com avanços importantes, mas ainda marcado por lacunas que limitam a profundidade analítica, a comparabilidade dos resultados e a reprodutibilidade das soluções propostas.

Os estudos analisados mostram predominância de abordagens baseadas em questionários, *surveys* e bases administrativas, além de crescimento no uso de dados de plataformas profissionais, especialmente o *LinkedIn*. Embora esse movimento amplie o potencial de análises em maior escala, a integração entre dados institucionais e dados da web ainda é incipiente, o que reduz a capacidade de construir interpretações mais robustas sobre as trajetórias profissionais.

No plano analítico, prevalece o uso de estatística descritiva, com adoção ainda limitada de técnicas de Aprendizado de Máquina e Mineração de Texto. De modo semelhante, as contribuições da Engenharia de Software permanecem pouco consolidadas, com recorrência de soluções *ad-hoc*, baixa documentação e escassa disponibilização de artefatos reutilizáveis, comprometendo replicabilidade e reuso.

A explicabilidade se destaca como a principal lacuna identificada. Em grande parte dos estudos, os processos analíticos e modelos são empregados sem mecanismos explícitos de interpretação, mantendo uma lógica de *caixa-preta* e restringindo a transparência e a confiança nos resultados. Esse cenário reforça a necessidade de abordagens que integrem múltiplas fontes de dados, adotem métodos analíticos mais avançados e incorporem explicabilidade como requisito.

Do ponto de vista prático, os achados deste mapeamento oferecem subsídios para gestores e formuladores de políticas educacionais estruturarem estratégias de acompanhamento de egressos mais transparentes e orientadas por evidências. Ao explicitar lacunas na integração de dados, na reprodutibilidade e no uso de explicabilidade, o estudo também apoia decisões institucionais relacionadas à avaliação curricular e ao monitoramento de trajetórias profissionais em perspectiva longitudinal.

Agradecimentos

As autoras agradecem à equipe do projeto Egress@s pelo apoio ao desenvolvimento desta pesquisa, especialmente nas discussões metodológicas e na organização dos dados. Agradecem também ao CNPq¹ e à FAPERGS² pelo apoio por meio de bolsas de pesquisa. Por fim, agradecem à UNIPAMPA³ pelo suporte institucional e pela infraestrutura disponibilizada.

¹Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

²Fundação de Amparo à Pesquisa do Estado do Rio Grande do Sul (FAPERGS).

³Universidade Federal do Pampa (UNIPAMPA).

Uso de Inteligência Artificial

Durante a elaboração deste artigo, foram utilizadas ferramentas de Inteligência Artificial Generativa para apoio à revisão linguística e organização textual. Essas ferramentas não foram empregadas para gerar resultados, analisar dados ou definir conclusões. Todas as decisões metodológicas e a responsabilidade pelo conteúdo permanecem integralmente com as autoras.

References

- [1] Adib Hakimi Abdul Rashid, Masurah Mohamad, Suraya Masrom, and Ali Selamat. Student career recommendation system using content-based filtering method. In *2022 3rd International Conference on Artificial Intelligence and Data Sciences (AiDAS)*, pages 60–65, 2022.
- [2] Libing Jiang. Research on market positioning of training for sports information-technology talents from the angle of employment direction. In *2010 3rd International Conference on Information Management, Innovation Management and Industrial Engineering*, volume 4, pages 463–467. IEEE, 2010.
- [3] Ken Chen, Qingnian Huang, Ming Hu, and Robert M. Randoy. Perspective: Employment conditions for computer engineering related college graduates in china. In *2011 6th International Conference on Computer Science & Education (ICCSE)*, pages 675–680, 2011.
- [4] Angeliq D. Lacasandile, Ronnet B. Altuna, Avonn C. Nova, Lilibeth P. Diaz, Kenneth R. Gunay, and Alezzandrei Ericka A. Balbuena. Fostering alumni involvement and professional advancement: A user-centric perspective on the national university alumni portal. In *2023 IEEE 15th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, pages 1–6, 2023.
- [5] Kais Dai, Celia González Nespereira, Ana Fernández Vilas, and Rebeca P Díaz Redondo. Scraping and clustering techniques for the characterization of linkedin profiles. *arXiv preprint arXiv:1505.00989*, 2015.
- [6] Sangita Lade, Ayush Billade, Akshat Chandrapatle, Srinivas Chenna, and Girish Chinchalpalle. Linkedin alumni profile data extraction. In *2024 4th International Conference on Pervasive Computing and Social Networking (ICPCSN)*, pages 174–178, 2024.
- [7] Rahmad Kurniawan, Siska Kurnia Gusti, Sukamto, Syamsudhuha, and Rahmad Abdillah. Predicting career trajectories of information system graduates using adaboost algorithm. In *2023 9th International Conference on Wireless and Telematics (ICWT)*, pages 1–5, 2023.
- [8] Zhong Minghui. Predicting graduates' career paths using neural networks and gaussian mixture models. In *2023 IEEE 6th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, pages 174–178, 2023.
- [9] Lucila Pena, Carla Curado, and Mírian Oliveira. The contribution of linkedin use to career outcome expectations. *Journal of Business Research*, 144:788–796, 2022.

- [10] Abdullah Hussein and Mohammed Hussain. Social-media based assessment of academic programs. *Studies in Educational Evaluation*, 62:149–157, 2019.
- [11] Lena Armstrong, Jayne Everson, and Amy J. Ko. Navigating a black box: Students' experiences and perceptions of automated hiring. In *Proceedings of the 2023 ACM Conference on International Computing Education Research - Volume 1*, ICER '23, page 148–158, New York, NY, USA, 2023. Association for Computing Machinery.
- [12] P.A. Kozlov, S.I. Safin, and V.O. Tolcheev. Identification and analysis of priority department topics of student's scientific research. In *2022 VI International Conference on Information Technologies in Engineering Education (Inforino)*, pages 1–4, 2022.
- [13] Lei Li, Guangzhi Zheng, Svetlana Peltsverger, and Chi Zhang. Career trajectory analysis of information technology alumni: A linkedin perspective. In *Proceedings of the 17th Annual Conference on Information Technology Education*, SIGITE '16, page 2–6, New York, NY, USA, 2016. Association for Computing Machinery.
- [14] S Abhilash, S Abhijith, GL Praveen, et al. Advanced career guidance system (acgs): With career road map. In *2025 International Conference on Networks & Advances in Computational Technologies (NetACT)*, pages 1–6. IEEE, 2025.
- [15] M. Ferreira, A. Barbosa, R. Braga, D. Saraiva, and C. Oliveira. Panorama comparativo da atuação acadêmica e profissional de egressas e egressos de um curso de bacharelado em ciência da computação. In *XXX Workshop sobre Educação em Computação*, pages 1–10, Porto Alegre, RS, Brasil, 2022. SBC.
- [16] Daphne Miedema, Toni Taipalus, Vangel V. Ajanovski, Abdussalam Alawini, Martin Goodfellow, Michael Liut, Svetlana Peltsverger, and Tiffany Young. Data systems education: Curriculum recommendations, course syllabi, and industry needs. In *2024 Working Group Reports on Innovation and Technology in Computer Science Education*, ITiCSE 2024, page 95–123, New York, NY, USA, 2025. Association for Computing Machinery.
- [17] Milena Soares Ferreira, Bianca Maia Ribeiro, Marcos Eduardo Lopes Silva, Camilla Borchhardt Quincozes, Alice Fonseca Finger, and Aline Vieira de Mello. Acompanhamento de egressos: Mapeamento sistemático de tópicos de interesse de cursos de computação. In *Anais do XXXV Simpósio Brasileiro de Informática na Educação (SBIE 2024)*, pages 2245–2258, Porto Alegre, RS, Brasil, 2024. SBC.
- [18] Barbara Kitchenham and Stuart Charters. Guidelines for performing systematic literature reviews in software engineering. EBSE Technical Report EBSE-2007-01, Software Engineering Group, School of Computer Science and Mathematics, Keele University and Department of Computer Science, University of Durham, July 2007. Version 2.3, 9 July 2007.
- [19] Kai Petersen, Robert Feldt, Shahid Mujtaba, and Michael Mattsson. Systematic mapping studies in software engineering. In *12th international conference on*

evaluation and assessment in software engineering (EASE). BCS Learning & Development, 2008.

- [20] Reinaldo V. Alvares, Marcelo S. Loutfi, and Nathielly de S. Campos. Onde estão meus egressos? relato sobre um mapeamento automatizado da vida profissional dos formados em sistemas de informação da unirio. *Anais do XXXI Workshop sobre Educação em Computação*, pages 1–10, 2020.
- [21] Maleerat Sodanil, Saranlita Chotirat, Lap Poomhiran, and Kanchana Viriyapant. Guideline for academic support of student career path using mining algorithm. In *Proceedings of the 2019 3rd International Conference on Natural Language Processing and Information Retrieval, NLPPIR '19*, page 133–137, New York, NY, USA, 2019. Association for Computing Machinery.
- [22] Henrique Soares Rodrigues and Reinaldo Viana Alvares. Em busca do perfil empreendedor dos egressos de mestrado e doutorado em computação no estado do rio de janeiro. In *Escola Regional de Sistemas de Informação do Rio de Janeiro (ERSI-RJ)*, pages 96–99. SBC, 2021.
- [23] Helen Ferreira, Edgar Oliveira, Reinaldo Braga, Mauro Oliveira, and Carina Oliveira. Um estudo do impacto de egressos no desenvolvimento regional como reflexo da interiorização do ensino superior. In *Anais do XXIX Workshop sobre Educação em Computação*, pages 428–437. Sociedade Brasileira de Computação, 2021.
- [24] Alice Finger, Andréa Bordin, and Aline de Mello. Perfil das egressas dos cursos de computação da unipampa: Uma análise da formação acadêmica e da atuação profissional. In *Anais do XIV Women in Information Technology*, pages 100–109, Porto Alegre, RS, Brasil, 2020. SBC.
- [25] Aline Vieira de Mello, Alice Fonseca Finger, and Andréa Sabedra Bordin. Ciência da computação e engenharia de software: semelhanças e diferenças a partir da realidade dos egressos. In *Anais do XXXI Simpósio Brasileiro de Informática na Educação (SBIE)*, pages 1773–1782. Sociedade Brasileira de Computação, 2020.
- [26] Venky Shankararaman and Swapna Gottipati. Mapping information systems student skills to industry skills framework. In *2016 IEEE Global Engineering Education Conference (EDUCON)*, pages 248–253, 2016.
- [27] Wantana Areeprayolkij and Mengzhen Li. A decision tree of bioengineering study and career path for educational guidance. In *2023 30th Asia-Pacific Software Engineering Conference (APSEC)*, pages 539–543. IEEE, 2023.
- [28] B. M. D. E Bannaka, D. M. H. S. G Dhanasekara, M. K Sheena, Anuradha Karunasena, and Nadeesa Pemadasa. Machine learning approach for predicting career suitability, career progression and attrition of it graduates. In *2021 21st International Conference on Advances in ICT for Emerging Regions (ICter)*, pages 42–48, 2021.
- [29] Pathathai Na Lumpoon and Pree Thiengburanathum. A design and implementation of performance dashboard for the work integrated learning unit. In *2019*

13th International Conference on Software, Knowledge, Information Management and Applications (SKIMA), pages 1–6. IEEE, 2019.

- [30] Travis D Breaux and Jennifer Moritz. A metric for measuring software engineering post-graduate outcomes. In *2023 IEEE/ACM 45th International Conference on Software Engineering: Software Engineering Education and Training (ICSE-SEET)*, pages 283–295. IEEE, 2023.
- [31] Theodoros Lappas. Mining career paths from large resume databases: Evidence from it professionals. *ACM Trans. Knowl. Discov. Data*, 14(3), May 2020.
- [32] Mohamad Jawhar, Majed Almotairi, Jeremy R Miller, and Shadi Jawhar. Ai-powered career matching system for university students: Bridging education and employment. In *2025 IEEE/ACIS 23rd International Conference on Software Engineering Research, Management and Applications (SERA)*, pages 113–117. IEEE, 2025.
- [33] Ewa Augustyniak and Piotr Augustyniak. Influencing the job market by the quality of graduates - a biomedical engineering example. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 3981–3984, 2015.
- [34] Linda Ott. Alumni as a resource to increase student retention in early computer science courses. In *2015 IEEE Frontiers in Education Conference (FIE)*, pages 1–9. IEEE, 2015.
- [35] Jiayi Guo, Lifen Song, Jialei Wang, Yuqi Wang, and Shiyu Zhao. An analysis of employment in agricultural and forestry colleges based on logistic modelling. In *2024 13th International Conference on Educational and Information Technology (ICEIT)*, pages 273–278. IEEE, 2024.
- [36] SL M Rukais, Saleem Adeeba, GAC A Herath, and BTG S Kumara. An ensemble model for predicting career paths of sri lankan it undergraduates. In *2025 International Research Conference on Smart Computing and Systems Engineering (SCSE)*, pages 1–6. IEEE, 2025.
- [37] Lara Gomes, Daniel Saraiva, Raquel Silveira, and Carina Oliveira. Um panorama comparativo por gênero em um curso de bacharelado em ciência da computação. In *Anais da XI Escola Regional de Computação do Ceará, Maranhão e Piauí*, pages 52–61, Porto Alegre, RS, Brasil, 2023. SBC.
- [38] R Ariyanto, C Rachmad, and AR Syulistyo. Automatically collect alumni data on social media. In *IOP Conference Series: Materials Science and Engineering*, volume 732, page 012071. IOP Publishing, 2020.
- [39] Adriana Postal, Clodis Boscarioli, and Jorge Bidarra. O papel da iniciação científica e do trabalho de conclusão de curso na formação acadêmica e profissional na Área da computação. In *Anais do XXI Workshop sobre Educação em Computação (WEI)*. Sociedade Brasileira de Computação, 2013.
- [40] Bharat Patel, Varun Kakuste, and Magdalini Eirinaki. Capar: a career path recommendation framework. In *2017 IEEE Third International Conference on Big Data Computing Service and Applications (BigDataService)*, pages 23–30. IEEE, 2017.

- [41] Thomas Louis Case, Hyo-Joo Han, and Eric Rimes. Career paths of computing program graduates: A linkedin analysis. In *Proceedings of the Southern Association for Information Systems Conference (SAIS)*, page 21, United States, 2016. Association for Information Systems.
- [42] Kais Dai, Celia González Nespereira, Ana Fernández Vilas, and Rebeca P. Díaz Redondo. Scraping and clustering techniques for the characterization of linkedin profiles. *CoRR*, abs/1505.00989, 2015.
- [43] Kundan Jha, D. Likhitha, M.Siri Chandana, M.Ram Prakash Reddy, and Maridu Bhargavi. Career prediction using machine learning. In *2024 8th International Conference on Inventive Systems and Control (ICISC)*, pages 118–122, 2024.
- [44] Mehrdad Maghsoudi. Uncovering the skillsets required in computer science jobs using social network analysis, 2023.
- [45] Ijeoma JF Ezika, Moses Olayemi, and Chibueze Tobias Orji. Career progression of electronic engineering graduates: A study of linkedin career profiles. In *2022 5th Information Technology for Education and Development (ITED)*, pages 1–6. IEEE, 2022.
- [46] Aniss Moumen, El Houcine Bouchama, and Younes EL BOUZEKRI EL IDIRISSI. Data mining techniques for employability: Systematic literature review. In *2020 IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS)*, pages 1–5. IEEE, 2020.
- [47] Neepa Kundu, Gourab Mandal, Anirban Bhandari, Udit Bhaskar, and Apurba Nandi. Deep q network-driven intelligent systems: Advancing the dynamics of student-alumni networking and professional mentorship. In *2025 International Conference on Next Generation Communication & Information Processing (INCIP)*, pages 445–450. IEEE, 2025.
- [48] Shanley Corvite and Julie Hui. Social media as a lens into careers during a changing world of work. *Proc. ACM Hum.-Comput. Interact.*, 8(CSCW2), November 2024.
- [49] Maria G Ioannides and Stylianos A Papazis. Development of new curricula in vocational education for the needs of enterprises and employability of graduates. In *2016 15th International Conference on Information Technology Based Higher Education and Training (ITHET)*, pages 1–5. IEEE, 2016.
- [50] M Sharmila Begum, JA Vishwa Raghawin, AM Mohamed Mubeen Azath, and A joel Amrith. Efficient way of interaction between alumni and students with the usage of sentiment analysis. In *2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC)*, pages 813–817. IEEE, 2022.
- [51] Mohamad Kassab. Navigating academic careers in software engineering: An empirical study of faculty and postdoctoral roles. In *Proceedings of the 33rd ACM International Conference on the Foundations of Software Engineering, FSE Companion '25*, page 936–946, New York, NY, USA, 2025. Association for Computing Machinery.

- [52] Priyank Kansal and Harsh Sadawarti. Empowering the future: An in-depth investigation of computer science graduate employability cognitive factors. In *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, pages 1–5. IEEE, 2023.