

Construção de uma ontologia para seleção de profissionais em TI

Wellynton Diniz¹, Adelma Andrade¹, Marcela Valença¹, César França¹

¹Cesar School
Recife – PE – Brasil

wjds@cesar.school, adelsantana@gmail.com

marcela.valenca@portodigital.org, franssa@cesar.school

Abstract. *The IT market is growing rapidly, but recruiting qualified professionals is challenging due to the lack of standardization in technical skills. A study developed an ontology to represent these competencies, using recruitment data from Porto Digital in Recife/PE. A matching algorithm was created to calculate the similarity between job vacancies and candidates, validating the ontology and the algorithm. Validation compared manual assessments by IT professionals with automatic assessments by the algorithm, demonstrating similar efficiency to traditional selection methods.*

Resumo. *O mercado de TI cresce rapidamente, mas o recrutamento de profissionais qualificados é desafiador devido à falta de padronização nas competências técnicas. O estudo desenvolveu uma ontologia para representar essas competências, usando dados de recrutamento do Porto Digital em Recife/PE. Um algoritmo foi criado para calcular a similaridade entre vagas e candidatos, validando a ontologia e o algoritmo. A validação comparou avaliações manuais de profissionais de TI com avaliações automáticas do algoritmo, mostrando eficiência semelhante aos métodos tradicionais de seleção.*

1. Introdução

A competitividade na indústria de software tem exigido profissionais dinâmicos e capacitados para acompanhar o ritmo do mercado. Na indústria de software, a busca por profissionais de TI com habilidades diversificadas é especialmente desafiadora. Alguns autores como [Ahmed et al. 2012, Gotel et al. 2009] relatam que o sucesso de um profissional de TI está ligado ao desenvolvimento de habilidades técnicas (*hard skills*) e habilidades não-técnicas (*soft skills*).

A falta de pessoas com as habilidades exigidas pelo mercado é conhecida como "skill gap" e se tornou um obstáculo no recrutamento de profissionais de TI [Agut and Grau 2002]. Estudos indicam que as deficiências de habilidades na área de TI estão relacionadas à formação acadêmica [Almi et al. 2011]. Além disto, a falta de precisão e uniformidade na definição de cargos na área de TI dificulta ainda mais o recrutamento de profissionais [Donohue and Power 2012].

A seleção de pessoal e os custos envolvidos demandam uma eficiência ao processo de recrutamento. A startup Gupy cita em matéria [Dias 2020] que "uma contratação errada pode custar 15 vezes um salário, segundo um levantamento realizado por Censo

Bazzola, diretor executivo da Bazz Consultoria em RH”. Isso porque há prejuízos com demissões, rotatividade de profissionais e treinamentos.

Diante disso, soluções tecnológicas baseadas em algoritmos têm sido desenvolvidas para melhorar o processo de recrutamento de pessoal. No entanto, a confiança nessas soluções ainda é limitada, pois os gestores preferem confiar em seus próprios julgamentos [Gandomani and Nafchi 2015]. Este artigo propõe uma abordagem baseada na “*design science*”, conforme descrito por [Dresch et al. 2015], para lidar com esses desafios no recrutamento de profissionais de TI. O objetivo principal foi construir uma ontologia de competências técnicas de TI para ser utilizada como base de conhecimento e desenvolver um algoritmo de similaridade para avaliá-la utilizando dados reais de um recrutamento realizado no Porto Digital. Foram identificadas necessidades de vagas das empresas e analisadas tecnologias e linguagens de programação representativas.

2. Fundamentação Teórica

2.1. Soluções tecnológicas para seleção de pessoas

A discrepância entre oferta e demanda de empregos cria desafios, como a necessidade de filtrar muitos candidatos. A inteligência artificial pode ser uma solução eficaz para esse problema [Sandoval 2020], entretanto, a resistência dos profissionais de RH em confiar totalmente nos algoritmos e a possibilidade de viés inconsciente nas ferramentas automatizadas são questões que precisam ser consideradas [Cappelli 2012].

Alternativas como sistemas de apoio à decisão baseados em mineração de dados têm se mostrado úteis para priorizar os esforços de recrutamento e aumentar a eficiência do processo [Mehta et al. 2013]. Assim, a padronização de linguagem entre empresas e candidatos, junto com o uso de tecnologias como ontologias, inteligência artificial e sistemas de apoio à decisão, surge como uma estratégia promissora para melhorar o recrutamento.

2.2. Ontologia como representação do conhecimento

O processo de representação do conhecimento é crucial para diversos campos científicos, permitindo a descrição clara de conceitos e termos específicos de um domínio [de Oliveira and de Carvalho 2008, Almeida et al. 2009]. As ontologias desempenham um papel fundamental nesse processo, fornecendo modelos formais de conhecimento que podem ser utilizados por ferramentas computacionais para apoiar a tomada de decisão [Castilho et al. 2008, Sewald Junior et al. 2013].

O autor [Gruber 1996] define ontologia como uma teoria lógica que oferece uma conceituação explícita, enquanto [Guarino 1998] classifica ontologias em tipos como Genéricas, de Domínio, de Tarefas e de Aplicação. No contexto de empresas internacionais, como Amazon e eBay, a criação de grandes estruturas taxonômicas é essencial para integrar conceitos de diferentes sistemas [Angermann and Ramzan 2017].

No campo da Engenharia de Software, ontologias têm sido utilizadas para compartilhar e reutilizar conhecimento durante o desenvolvimento de sistemas [Cota et al. 2004, Holsapple and Joshi 2002] enfatizam que as ontologias podem automatizar a comunicação entre pessoas, processos, organizações e sistemas, tornando-se essenciais para a expansão da automatização nos diversos campos.

3. Solução proposta

3.1. Método de geração de uma ontologia

Os princípios que deram sustentação científica ao processo de construção da ontologia foram fundamentados a partir da metodologia proposta por [Angermann and Ramzan 2017], com a construção de taxonomias criando-se hierarquias em forma de árvores e da metodologia 101 [Noy and McGuinness 2001], que sugerem uma modelagem simples e intuitiva, com avaliações e depurações contínuas durante todo o ciclo de vida da ontologia.

3.2. Estruturas de referência

O *Rational Unified Process* – RUP – (1987; 2001) propôs um modelo de *job description* estabelecendo um conjunto de conceitos e definições de papéis dos profissionais de engenharia de software e CBO define as ocupações e suas responsabilidades de formas genéricas e organiza ocupações em famílias, cada família tem detalhes específicos definidos pelo CBO, incluindo formação, experiência, atividades e qualificações necessárias.

Conforme [Donohue and Power 2012], a falta de clareza na definição de cargos gera interpretações diversas entre as empresas. As abordagens do CBO e RUP revelam descrições limitadas e genéricas dos cargos. Baseando-se nisso, conforme sugerido por [Chin and Chang 2009], a ontologia proposta foi desenvolvida partindo-se de insumos que foram sendo incorporados gradativamente durante a evolução desta pesquisa.

3.3. Proposta de ontologia sobre conhecimentos técnicos de TI

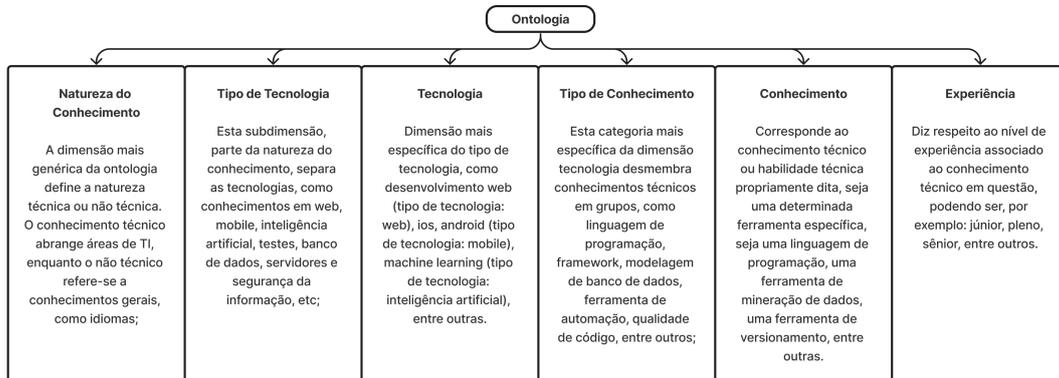
Foi desenvolvida uma ontologia para padronizar as habilidades técnicas de empresas e candidatos de TI. Isso inclui a criação de taxonomias de competências e habilidades, utilizadas para desenvolver um algoritmo de correspondência entre vagas de emprego e candidatos qualificados. A construção da ontologia seguiu a metodologia 101 (Figura 1).

Figura 1. Processo de construção da ontologia

Etapa 1 Escopo da ontologia:	Etapa 2 Verificação de ontologias existentes:	Etapa 3 Enumeração de termos importantes na ontologia:	Etapa 4 Definição das classes e hierarquias:	Etapa 5 Definição das propriedades das classes:	Etapa 6 Definição das restrições das classes:	Etapa 7 Criação de instâncias:
Determinado pela análise dos processos seletivos do Porto Digital e de fontes como portais de recrutamento online, Carrer Smart, CBO e RUP. O objetivo foi estabelecer uma linguagem comum para as competências técnicas de TI.	Foi proposta uma heurística para organizar os dados e estabelecer caminhos comuns. Não havia uma ontologia relevante preexistente, então o desenvolvimento ocorreu informalmente, do zero.	Foram mapeados os termos essenciais exigidos pelas vagas e apresentados pelos candidatos de TI, removendo conectivos e termos acessórios para manter apenas substantivos, excluindo artigos, advérbios, adjetivos e verbos....	Durante a análise dos dados, foram definidas relações entre classes, como natureza de conhecimento, linguagem de programação e experiência. Uma estrutura de árvore foi então criada com um número fixo de níveis.	Foram definidas através da construção de um modelo de banco de dados relacional, contendo a representação da ontologia e de seus elementos através de tabelas	Foram definidas a partir de consultas nas tabelas do banco de dados, sendo possível incluir filtros para seleção de campos e operações lógicas com os dados inseridos.	A criação das instâncias pôde ser realizada a partir do procedimento de carga na estrutura das tabelas do banco de dados.

Fundamentando-se no estudo de [Guarino 1998], onde termos da linguagem da ontologia se encontram relacionados a conceitos semanticamente compartilhados, formando redes interligadas, foi desenvolvido o modelo ontológico com a estrutura de domínio baseada nos conhecimentos técnicos dos profissionais de TI. No contexto deste estudo, conforme relevância do estudo de [Angermann and Ramzan 2017], a ontologia contém as superclasses e subclasses, que serão nomeadas como dimensões (Figura 2).

Figura 2. Ontologia em dimensões



Os cargos foram agrupados com base em responsabilidades, requisitos e qualificações semelhantes, assim como os perfis exigidos pelas vagas e informados pelos candidatos, contextualizando suas áreas de atuação.

4. Metodologia de validação experimental

Esta seção descreve a metodologia empregada para avaliar a solução proposta. Inclui o modelo de pesquisa, a avaliação manual, a avaliação automática e uma análise comparativa entre elas. Este processo está alinhado com a abordagem proposta pela *design science research* [Dresch et al. 2015].

4.1. Modelo de pesquisa

Na fase de avaliação do artefato, modelos estruturados foram criados com dados de vagas e currículos para validar a ontologia e o algoritmo proposto. Os candidatos foram classificados manualmente e o algoritmo de similaridade foi implementado separadamente para processar os dados. Os resultados foram comparados usando testes estatísticos.

4.2. Avaliação manual

O processo de avaliação manual foi realizado seguindo 6 etapas (Figura 3).

Figura 3. Etapas do processo de avaliação manual

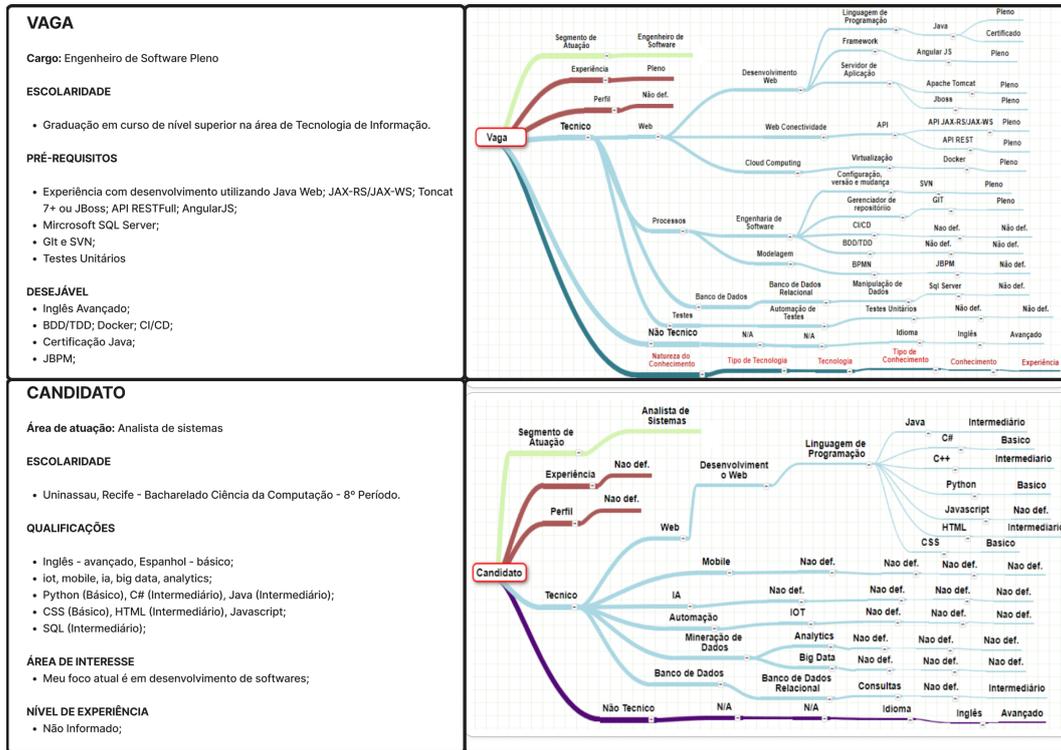


Etapa 1: Comparação do algoritmo proposto com dados reais de recrutamento do Porto Digital, classificação de candidatos em ordem de adequação às vagas, usando dois modelos (Real e Ontologia).

Etapa 2: No modelo real, as vagas e candidatos são representados de forma tradicional, enquanto no modelo de ontologia, são considerados pré-requisitos e conhecimentos técnicos desejados.

Etapa 3: A figura 4 apresenta a estrutura de dados da vaga e do candidato no modelo real e no modelo da ontologia.

Figura 4. Apresentação da vaga no modelo real X modelo ontologia



Etapa 4: Foram escolhidos 4 vagas e 20 candidatos aleatoriamente, eles foram divididos aleatoriamente em 4 grupos, cada um com 1 vaga e 5 candidatos.

Etapa 5: Participaram 12 avaliadores, cada um atribuído a um grupo específico para garantir análises abrangentes. Foram selecionados com base em experiência em TI e em processos de recrutamento.

Etapa 6: Os avaliadores foram informados sobre o objetivo da pesquisa, seguido pela análise das vagas e dos candidatos nos dois modelos. Além disso, atribuíram importância aos conhecimentos técnicos numa escala de 1 a 5. Em seguida, preencheram o formulário comparativo via *Google Forms*. A ordem de apresentação dos modelos foi invertida em cada execução para evitar viés, seguindo um delineamento denominado "quadrado latino", adequado para ambientes experimentais heterogêneos.

4.3. Avaliação automática

O processo de avaliação automática foi realizado seguindo 3 etapas, sendo elas: Proposição do Algoritmo; Implementação do Algoritmo; Processamento de dados;

Na Etapa 1, o algoritmo proposto visa aprimorar a correspondência entre vagas e candidatos, comparando habilidades exigidas pelas vagas com as declaradas pelos candidatos, de forma similar à busca por diagnósticos na CID [Guarino 1998, Cota et al. 2004, Holsapple and Joshi 2002]. Fundamentado na ontologia de domínio, o algoritmo hierarquiza as competências requeridas pelas vagas e as habilidades dos candidatos, permitindo uma comparação eficiente entre as árvores hierárquicas, reduzindo o tempo computacional e o viés subjetivo. Ele estabelece o grau de correspondência entre as estruturas de domínio, definindo um escore de similaridade total ou parcial,

além de considerar o valor de importância atribuído pelo recrutador a cada habilidade [Mehta et al. 2013, Malinowski et al. 2006].

Na Etapa 2, para implementar o algoritmo, foi utilizado o banco de dados *MySQL*. Foram criadas entidades relacionadas a vagas, candidatos e uma entidade para relacionar vagas e candidatos para armazenar os escores de similaridade. Além disso, foi desenvolvida uma *stored procedure* para calcular e armazenar os escores de similaridade.

Na Etapa 3, o processamento dos dados resultou nos escores de similaridade entre vagas e candidatos, considerando os pesos atribuídos pelos profissionais de TI. Após a conclusão das etapas de implementação, os dados foram processados e os escores dos candidatos para as vagas foram coletados.

4.4. Método de análise comparativa

Para análise, foi criada uma matriz para classificar os candidatos, separando as avaliações por vaga, candidato e avaliador. Isso possibilitou a comparação entre a classificação do modelo real e do modelo de ontologia (avaliação manual) com a classificação gerada pelo algoritmo, com e sem a influência do valor de importância (avaliação automática).

O método estatístico Mann-Whitney U Test foi escolhido para avaliar as discrepâncias entre as avaliações, levando em consideração o cenário e as limitações do algoritmo. Este teste não paramétrico é adequado para comparar e ordenar grupos de dados ordinais, sem fazer suposições sobre a distribuição dos dados.

5. Resultados

5.1. Resultado das avaliações manuais

A avaliação envolveu 12 avaliadores, 4 vagas e 20 candidatos distribuídos em 4 grupos. A análise utilizou como base dados do processo seletivo "Emprego" do Porto Digital. Os avaliadores selecionados, em sua maioria, possuíam mais de 15 anos de experiência.

A avaliação manual foi baseada em perguntas objetivas e discursivas. As objetivas são acompanhadas de justificativas dos participantes, enquanto as discursivas visam a avaliação geral e aprimoramento do modelo em estudo. As perguntas podem ser visualizadas no repositório do zenodo (<https://zenodo.org/records/10914411>).

Em relação a clareza das vagas e dos candidatos. No modelo real as vagas estavam bem definidas, mas sugeriram mais detalhes nos conhecimentos técnicos. Quanto aos candidatos, os currículos estavam pouco claros. No modelo de ontologia, os avaliadores concordaram que as informações sobre as vagas estavam claras, mas sugeriram mais detalhes nas habilidades. Quanto aos candidatos, a maioria achou claras as descrições do segmento de atuação e das habilidades técnicas, mas pediram mais detalhes.

A análise do tempo revelou que 6 dos 12 avaliadores gastaram menos tempo com o modelo de ontologia para avaliar candidatos, enquanto 5 gastaram menos tempo com o modelo real. Apenas 1 avaliador teve tempos iguais em ambos. No geral, o modelo real levou menos tempo, mas a diferença não foi significativa.

Em relação aos critérios de classificação dos candidatos, a compatibilidade das qualificações técnicas com as exigências da vaga foi destacada por 11 avaliadores como o fator mais relevante, fortalecendo a base do estudo. O nível de experiência foi mencionado

por 3 avaliadores, seguido pela formação acadêmica e área de interesse, ambos com 2 menções. A atuação do candidato recebeu 1 menção, enquanto o conhecimento em inglês foi considerado relevante por um avaliador.

O resultado que os avaliadores chegaram é apresentado na figura 5. Os candidatos são classificados de 1 a 5, sendo o 1 o mais provável para entrevista e o 5 o menos provável para entrevista no processo de seleção.

Figura 5. Resultado da avaliação manual

Vaga 1	Avaliação Manual		Vaga 2	Avaliação Manual		Vaga 3	Avaliação Manual		Vaga 4	Avaliação Manual	
	Modelo Real	Modelo Ontologia									
Avaliador 1			Avaliador 4			Avaliador 7			Avaliador 10		
Candidato 4	5	5	Candidato 17	3	5	Candidato 16	5	5	Candidato 11	1	1
Candidato 8	3	3	Candidato 19	5	4	Candidato 20	4	3	Candidato 18	5	2
Candidato 9	2	1	Candidato 24	4	3	Candidato 21	1	1	Candidato 23	2	5
Candidato 13	4	2	Candidato 27	1	2	Candidato 25	2	2	Candidato 26	4	4
Candidato 14	1	4	Candidato 28	2	1	Candidato 30	3	4	Candidato 29	3	3
Avaliador 2			Avaliador 5			Avaliador 8			Avaliador 11		
Candidato 4	5	4	Candidato 17	4	5	Candidato 16	5	5	Candidato 11	2	4
Candidato 8	3	3	Candidato 19	5	4	Candidato 20	1	2	Candidato 18	5	3
Candidato 9	2	1	Candidato 24	1	3	Candidato 21	2	3	Candidato 23	3	5
Candidato 13	1	2	Candidato 27	3	2	Candidato 25	4	4	Candidato 26	1	1
Candidato 14	4	5	Candidato 28	2	1	Candidato 30	3	1	Candidato 29	4	2
Avaliador 3			Avaliador 6			Avaliador 9			Avaliador 12		
Candidato 4	5	5	Candidato 17	5	5	Candidato 16	4	5	Candidato 11	1	1
Candidato 8	3	3	Candidato 19	4	4	Candidato 20	5	4	Candidato 18	5	3
Candidato 9	2	1	Candidato 24	3	2	Candidato 21	2	2	Candidato 23	4	2
Candidato 13	1	2	Candidato 27	1	1	Candidato 25	1	1	Candidato 26	3	4
Candidato 14	4	4	Candidato 28	2	3	Candidato 30	3	3	Candidato 29	2	5

Para a Vaga 1: As classificações dos candidatos variaram levemente entre os avaliadores. O Avaliador 1 teve mudanças notáveis entre os modelos real e de ontologia, influenciado pelo conhecimento em redes e suporte técnico. O Avaliador 2 manteve principalmente as mesmas posições, com exceção do Candidato 8. Já o Avaliador 3 alternou entre os Candidatos 9 e 13 nos primeiros lugares, com os demais permanecendo estáveis. Em geral, as avaliações não mostraram diferenças significativas entre os avaliadores.

Para a Vaga 2: Avaliadores notaram variações nas classificações dos candidatos. Destaque para o Candidato 17, com maior deslocamento do 3º para o 5º lugar, e o candidato 24, do 1º para o 3º lugar. Os avaliadores apontaram falta de detalhes sobre a experiência dos candidatos com as tecnologias apresentadas, prejudicando a análise curricular. No geral, as avaliações mostraram disparidades, possivelmente devido à inconsistência nos dados curriculares.

Para a Vaga 3: Avaliadores notaram variações nas classificações. Avaliador 7 destacou dois candidatos empatados. Avaliador 8 mencionou variações significativas em algumas classificações, baseando-se nas tecnologias exigidas. Avaliador 9 observou poucas mudanças nas classificações, porém achou o modelo de ontologia confuso. No geral, as avaliações mostraram consistência, apesar de algumas variações entre os modelos.

Para a Vaga 4: Os avaliadores classificaram os candidatos com base em critérios como qualificação, área de interesse e experiência, com algumas inversões entre os modelos. Avaliador 10 destacou que a reclassificação de um candidato se deveu ao conhecimento em tecnologias web. Avaliadores 11 e 12 notaram discrepâncias na reclassificação de um candidato entre os modelos, atribuindo isso à percepção das qualificações. No geral, as avaliações mostraram semelhanças, resultando em diferentes classificações.

Durante a classificação, houveram empates entre candidatos. Os avaliadores usa-

ram critérios como quantidade e conhecimento de tecnologias, área de interesse e experiência para desempatar. Esses critérios refletem a importância de habilidades e experiências diversas, conforme evidenciado em estudos anteriores.

Dos 12 avaliadores, 10 veem potencial no modelo de ontologia para auxiliar no recrutamento, mas ressaltam a necessidade de melhor exploração e estruturação da ontologia e a abordagem de aspectos comportamentais dos candidatos. Enquanto 2 veem apenas um suporte no processo de recrutamento e seleção.

5.2. Resultado das avaliações automáticas

Durante a avaliação manual, os avaliadores atribuíram valores de importância a cada linha de conhecimento técnico necessário para a vaga. Isso permitiu analisar a relevância de cada habilidade técnica em relação às outras, conforme a percepção de cada avaliador.

Na avaliação automática, a correspondência entre a linha da vaga e a do candidato foi essencial. Essa métrica afeta o resultado apenas com uma correspondência total de 100% ou parcial de 80%. Os dados resultantes incluíram três níveis de importância: 1 para muito importante (adicionando 10 pontos ao escore do candidato), 3 para importante (adicionando 5 pontos) e 5 para pouco importante (acrescentando 1 ponto).

Os candidatos foram ordenados de acordo com a similaridade com a vaga, sendo o mais semelhante classificado como 1º e o menos semelhante como 5º. Se os candidatos tiverem escores iguais, a ordenação foi determinada pela identificação do candidato. Por exemplo, se dois candidatos tiverem escores iguais, o candidato 1 ocupará a posição 1 e o candidato 2 ocupará a posição 2, mesmo com escores iguais.

Para a vaga 1, observa-se que os candidatos não tiveram suas classificações alteradas pelos avaliadores independentemente do valor de importância. O candidato 4 não atendeu aos requisitos da vaga (total e parcial), ao contrário dos outros candidatos. Além disso, o valor de importância atribuído pelos avaliadores resultou em escores semelhantes, indicando uma mudança proporcional, porém não significativa.

Para a vaga 2, observamos um resultado semelhante, onde as posições dos candidatos não foram alteradas entre os avaliadores, independentemente do valor de importância. Todos mostraram uma relação de similaridade total ou parcial com os requisitos da vaga. Embora tenha havido variação nos escores devido aos diferentes pesos atribuídos pelos avaliadores, isso não causou uma mudança significativa na classificação.

Para a vaga 3, observamos uma conclusão semelhante, onde as posições dos candidatos permaneceram inalteradas entre os avaliadores, independentemente do valor de importância. Todos esses candidatos mostraram uma relação de similaridade total ou parcial com os requisitos da vaga. Embora tenha havido variação nos escores devido aos diferentes pesos atribuídos pelos avaliadores, isso não causou uma mudança significativa.

Para a vaga 4, ao analisar a classificação e o escore dos candidatos na avaliação automática, notamos que os candidatos obtiveram escores idênticos, indicando um empate técnico. No entanto, entre os avaliadores 10, 11 e 12, houve uma mudança nos escores dos candidatos quando o valor de importância foi aplicado pelo avaliador 12.

Podemos visualizar com melhor clareza tal classificação através do candidato 26, por exemplo, que teve uma pontuação maior (220 pontos), seguido dos candidatos 11, 18

e 23 com a mesma pontuação (215 pontos) e por último o 29 (110 pontos). Ao aplicar o valor de importância dado pelo avaliador 12, a similaridade entre a vaga e o candidato 26 aumentou, fazendo com que o mesmo fosse reclassificado para a posição 1.

Tendo em vista a percepção entre os avaliadores, pode-se afirmar que, diante dos contextos apresentados, o valor de importância é impactante na ordenação dos candidatos, justificando assim sua aplicabilidade no cálculo do escore de similaridade.

5.3. Análise comparativa das avaliações

A análise compara a avaliação manual com a avaliação automática, buscando identificar os principais motivos de semelhanças ou diferenças entre as duas avaliações. Esse processo foi conduzido utilizando o teste não paramétrico Mann-Whitney U Test, utilizando o site *Social Science Statistics* como ferramenta de apoio. O valor atribuído representa a posição do candidato na classificação de um determinado grupo vaga x candidatos.

Para fins avaliativos, parâmetros comparativos foram abordados durante o processo. São eles: Modelo Real e Modelo Ontologia; Modelo Real e Algoritmo Sem Valor de Importância; Modelo Real e Algoritmo Com Valor de Importância; Modelo Ontologia e Algoritmo Sem Valor de Importância; Modelo Ontologia e Algoritmo Com Valor de Importância; Algoritmo Sem Valor de Importância e Algoritmo Com Valor de Importância;

Primeiramente foi realizada avaliação comparativa entre os modelos real e ontologia. Tal avaliação não resultou em diferenças significativas (**U=3421, Z=- 0.0798, P=.93624**). O mesmo foi observado para a comparação entre as avaliações do modelo real e do algoritmo sem a incidência do valor de importância do avaliador (**U=3421, Z=-0.2063, P=.83366**). Considerando o modelo real e o algoritmo com o valor de importância temos (**U=3421, Z=-0.2501, P=.80258**), que não resultou em diferenças significativas.

Em se tratando da comparação do modelo de ontologia e do algoritmo sem incidência do valor de importância do avaliador temos um resultado sem diferença significativa (**U=3421, Z=-0.3045, P=.76418**), o que também observamos na análise modelo de ontologia e do algoritmo com incidência do avaliador (**U=3421, Z=-0.1954, P=.84148**).

Em relação à análise comparativa das amostras do algoritmo sem e com a incidência do valor de importância, foi evidenciado que o tamanho da amostra é muito pequeno para permitir um cálculo confiável, isso porque o teste despreza as amostras idênticas entre os pares. A partir de um maior volume de dados, o cálculo poderia ser representativo para as hipóteses em estudo.

Resumidamente, a análise dos dados e os critérios definidos para o valor de importância no algoritmo de correspondência indicam que, embora a importância atribuída pelos avaliadores não tenha tido um impacto significativo na ordem dos candidatos e a análise estatística não tenha mostrado diferenças significativas, ainda assim foram observadas algumas disparidades entre as avaliações.

Na Vaga 1, os candidatos 9, 13 e 8 obtiveram os maiores escores na avaliação automática. No modelo de ontologia, o avaliador 1 reclassificou o candidato 13 para o 2º lugar, afirmando que a estrutura facilitou a avaliação. Essa mudança resultou em uma ordenação igual à avaliação automática para o candidato 13. Quanto aos outros candidatos, apenas uma posição foi alterada para o candidato 4, indo do 5º para o 4º lugar.

Na vaga 2, os candidatos 27, 28 e 24 obtiveram os maiores escores na avaliação automática. O avaliador 4 reclassificou o candidato 17 da posição 3 para 5 entre os modelos real e de ontologia, resultando em uma posição mais próxima à avaliação automática (posição 5 e 4), devido à similaridade do escore com os candidatos classificados abaixo. Os demais candidatos tiveram deslocamentos mínimos de apenas 1 posição. Isso sugere uma análise mais coerente com as habilidades e conhecimentos declarados pelos candidatos, dadas as exigências diversificadas da vaga (desenvolvedor *fullstack*), o que justificaria a argumentação do avaliador em relação à complexidade e nível de detalhamento da vaga.

Na vaga 3, os candidatos 21, 30 e 20 obtiveram os maiores escores. O avaliador 7 reclassificou o candidato 25 da posição 2 para 5 na avaliação do algoritmo, devido ao seu escore mais baixo em relação aos outros, indicando menor similaridade com a vaga. Considerando os candidatos 20 e 30 "tecnicamente empatados", conforme semelhança de seus escores (280 e 300 pontos, respectivamente) e na inversão de suas classificações.

Na vaga 4, os candidatos 11, 18, 23 e 26 obtiveram a mesma pontuação de 200 avaliação automática. No entanto, o algoritmo não inclui critérios para desempatar candidatos com pontuações iguais. Assim, a similaridade entre os candidatos e os requisitos da vaga resultou na mesma pontuação para todos. Para resolver esse empate, é necessário atribuir um peso de importância aos critérios de avaliação.

6. Conclusões

Encontrar profissionais de TI com competências diversificadas para atender às necessidades das empresas é uma tarefa difícil. Por mais que as empresas façam uso de alternativas automatizadas para realizar o recrutamento, muitos recrutadores mencionam relações de *match* falhas, passíveis de propensões preconceituosas.

Os Resultados enfatizaram a importância da compatibilidade entre habilidades técnicas exigidas e as dos candidatos. Descrições de vagas confusas e currículos carentes de detalhes prejudicam o processo seletivo. Essas questões inicialmente limitaram a compreensão da ontologia, mas sua validade para análise curricular foi confirmada.

A pesquisa mostrou que o modelo ontológico pode ser usado para correspondências entre vagas e candidatos, similar aos métodos tradicionais de recrutamento, confirmado pela comparação entre o modelo real e o modelo de ontologia, tanto manual quanto automaticamente. A comparação entre essas avaliações não revelou diferenças significativas, apesar das particularidades e critérios subjetivos de cada método.

No entanto, o estudo possui limitações. As conclusões são válidas, mas podem mudar com maior quantidade de dados. Além disso, as análises são baseadas nos 12 avaliadores, profissionais de TI, embora experientes, não são especialistas em RH, o que pode não representar outras realidades. Ainda, o detalhamento das vagas e dos candidatos foi subjetivo, com falta de especificação adequada de conhecimentos técnicos. A falta de consideração dos aspectos comportamentais dos candidatos na estrutura da ontologia. Essas ameaças não impedem a generalização dos resultados, mas, como são preliminares, não são conclusivos, apenas indicativos.

Para trabalhos futuros, sugere-se criar uma ferramenta automatizada baseada em ontologia de domínio para preencher dados de vagas e currículos, estabelecendo uma estrutura padronizada.

Referências

- Agut, S. and Grau, R. (2002). Managerial competency needs and training requests: The case of the spanish tourist industry. *Human Resource Development Quarterly*, 13(1):31–52.
- Ahmed, F., Capretz, L. F., and Campbell, P. (2012). Evaluating the demand for soft skills in software development. *It Professional*, 14(1):44–49.
- Almeida, M. B., Souza, R. R., and Bax, M. P. (2009). O grupo de pesquisa representação do conhecimento, ontologias e linguagem: atividades e perspectivas. In *Anais do 2º Seminário de Pesquisa em Ontologia no Brasil*, Rio de Janeiro. Editora da UFMG.
- Almi, N. E. A. M., Rahman, N. A., Purusothaman, D., and Sulaiman, S. (2011). Software engineering education: The gap between industry’s requirements and graduates’ readiness. In *2011 IEEE Symposium on Computers & Informatics*, pages 542–547. IEEE.
- Angermann, H. and Ramzan, N. (2017). *Taxonomy Matching Using Background Knowledge: Semantic Web and Heterogeneous Repositories*. Springer International Publishing, Suíça.
- Cappelli, P. (2012). *Why Good People Can’t Get Jobs: The Skills Gap and What Companies Can Do About It*. Kindle Edition, Filadélfia.
- Castilho, L. V., Lopes, H. S., and Tacla, C. A. (2008). Modeling and building an ontology for neuropediatric physiotherapy domain. In *Proceedings of the 8th International Conference on Hybrid Intelligent Systems*, volume 1, pages 210–215, Barcelona; Piscataway. IEEE Computer Press.
- Chin, K. L. and Chang, E. (2009). Sustainable ict education ecosystem. In *2009 3rd IEEE International Conference on Digital Ecosystems and Technologies*, pages 271–277. IEEE.
- Cota, R. I., Menezes, C., and Falbo, R. A. (2004). Modelagem organizacional utilizando ontologias e padrões de análise. Technical report, NEMO (Núcleo de Estudos em Modelagem Conceitual e Ontologias), Vitória.
- de Oliveira, H. C. and de Carvalho, C. L. (2008). Gestão e representação do conhecimento. *São Paulo: UFG*.
- Dias, G. (2020). 4 custos exorbitantes da má contratação.
- Donohue, P. and Power, N. (2012). Legacy job titles in it: the search for clarity. In *Proceedings of the 50th annual conference on Computers and People Research*, pages 5–10.
- Dresch, A., Lacerda, D. P., and Junior, J. A. V. A. (2015). *Design science research: método de pesquisa para avanço da ciência e tecnologia*. Bookman Editora.
- Gandomani, T. J. and Nafchi, M. Z. (2015). An empirically-developed framework for agile transition and adoption: A grounded theory approach. *Journal of Systems and Software*, 107:204–219.
- Gotel, O., Kulkarni, V., Say, M., Scharff, C., and Sunetnanta, T. (2009). A global and competition-based model for fostering technical and soft skills in software engineering

- education. In *2009 22nd Conference on Software Engineering Education and Training*, pages 271–278. IEEE.
- Gruber, T. (1996). What is an ontology? *Knowledge Systems*.
- Guarino, N. (1998). *Formal Ontology and Information Systems*. IOS Press, Amsterdam; Países Baixos.
- Holsapple, C. W. and Joshi, K. (2002). A collaborative approach to ontology design. *Commun. ACM*, 45(2):42–47.
- Malinowski, J., Keim, T., Wendt, O., and Weitzel, T. (2006). Matching people and jobs: a bilateral recommendation approach. In *Proceedings of the 39th Hawaii International Conference on System Sciences*, page 137c, Kauia, EUA. IEEE.
- Mehta, S., Pimplikar, R., Singh, A., Varshney, L., and Visweswariah, K. (2013). Efficient multifaceted screening of job applicants. In *Proceedings of the 16th International Conference on Extending Database Technology*, pages 661–671, Genova. EDBT.
- Noy, N. F. and McGuinness, D. (2001). Ontology development 101: A guide to creating your first ontology.
- Sandoval, G. (2020). Seleção esperta. *Exame*, 54(3):50–53.
- Sewald Junior, E., Rotta, M., Silva, E., Rover, A., Gauthier, F., and Carmo, V. (2013). Processo de desenvolvimento de ontologia de domínio da área jurídica: caso acidentes de trânsito. In *Actas da 8ª Conferência Ibérica de Sistemas e Tecnologias de Informação*, volume 2, pages 212–217, Lisboa. Associação Ibérica de Sistemas e Tecnologias de Informação.