

AS COMPETÊNCIAS DOS PROFISSIONAIS DA INFORMAÇÃO PARA A ANÁLISE E ORGANIZAÇÃO DE DADOS

proposta conceitual e aplicação da OpenAI em anúncios de emprego na web

Francisco-Carlos Paletta¹

Universidade de São Paulo
fcpaletta@usp.br

Juan-Antonio Pastor-Sánchez²

Universidad de Murcia
pastor@um.es

Barbara Coelho Neves³

Universidade Federal da Bahia
barbaran@ufba.br

José-Antonio Moreira-González⁴

Universidad Carlos Tercero de Madrid
jamore@bib.uc3m.es

Resumo

Se apresenta uma proposta conceitual e metodológica sobre as competências dos profissionais da informação (PI) na análise e organização de dados, a partir de ofertas de emprego divulgadas na web. Parte-se da premissa de que os dados são fundamentais para a tomada de decisões em contextos empresariais e institucionais, e que os PI, historicamente capacitados em mediação informacional, organização do conhecimento e uso de TIC, têm se inserido em equipes interdisciplinares voltadas à análise de dados. A pesquisa analisa 84 anúncios de emprego extraídos de portais internacionais (ALAJobLIST, Glassdoor, Catho, LinkedIn), cujos requisitos envolvem análise de dados, uso de taxonomias ou outros sistemas de organização do conhecimento (SOC), e formação em Ciência da Informação. As competências extraídas foram processadas com apoio da OpenAI e organizadas em oito macrogrupos: gestão empresarial, gestão de dados, organização da informação, análise estatística, programação, tecnologias web, capacidade analítica e temas específicos. Os resultados indicam que as empresas valorizam o domínio técnico e interpessoal dos PI, além de sua familiaridade com padrões, softwares e representações semânticas. Contudo, exigem também formação sólida em estatística e capacidade de liderança. Conclui-se que o PI possui papel relevante e competitivo no mercado de dados, demandando constante atualização profissional.

Palavras-chave: ciência da informação; profissional da informação; mercado de trabalho; anúncios de emprego na web; competências; experiência; análise de conteúdo; openai; análise de dados; organização taxonômica.

¹ Professor Pesquisador da Universidade de São Paulo, Escola de Comunicações e Artes, Brasil.

² Profesor Pesquisador da Facultad de Comunicación y Documentación, Espanha.

³ Professora Pesquisadora da Universidade Federal da Bahia (UFBA), Programa de Pós-Graduação em Ciência da Informação, Brasil

⁴ Professor Visitante na Universidade Federal da Bahia (UFBA). Programa de Pós-Graduação em Ciência da Informação Univ. Pesquisador da Carlos III de Madrid, Depto. de Biblioteconomía y Documentación, Espanha



Esta obra está licenciada sob uma licença

Creative Commons Attribution 4.0 International (CC BY-NC-SA 4.0).

P2P & INOVAÇÃO, Rio de Janeiro, v. 11, n. 2, p. 1-42, e-7640, jan./jun. 2025.

INFORMATION PROFESSIONALS COMPETENCIES FOR DATA ANALYSIS AND ORGANIZATION

a conceptual framework and an OpenAI-based analysis of online job ads

Abstract

This paper presents a conceptual and methodological proposal regarding the competencies of information professionals (IPs) in data analysis and organization, based on job advertisements published on the web. It starts from the premise that data is fundamental for decision-making in business and institutional contexts, and that IPs—historically trained in information mediation, knowledge organization, and the use of ICTs—have increasingly become part of interdisciplinary data analysis teams. The study analyzes 84 job postings collected from international platforms (ALAJobLIST, Glassdoor, Catho, LinkedIn), whose requirements involve data analysis, use of taxonomies or other knowledge organization systems (KOS), and a background in Information Science. The competencies extracted were processed with the support of OpenAI and organized into eight macro-groups: business management, data management, information organization, statistical analysis, programming, web technologies, analytical skills, and specific topics. The results indicate that companies value both the technical and interpersonal skills of IPs, as well as their familiarity with standards, software tools, and semantic representations. However, they also demand solid training in statistics and leadership capacity. The study concludes that IPs hold a relevant and competitive role in the data market, requiring continuous professional development.

Keywords: information science; information professional; labor market; web job ads; competencies.

COMPETENCIAS DE LOS PROFESIONALES DE LA INFORMACIÓN PARA EL ANÁLISIS Y LA ORGANIZACIÓN DE DATOS

propuesta conceptual y aplicación de OpenAI en anuncios de empleo en la web

Resumen

Se presenta una propuesta conceptual y metodológica sobre las competencias de los profesionales de la información (PI) en análisis y organización de datos, basada en ofertas de empleo publicadas en la web. Se parte de la premisa de que los datos son fundamentales para la toma de decisiones en contextos empresariales e institucionales, y que los PI, con una formación histórica en mediación de información, organización del conocimiento y uso de las TIC, se han integrado en equipos interdisciplinarios centrados en el análisis de datos. La investigación analiza 84 anuncios de empleo extraídos de portales internacionales (ALAJobLIST, Glassdoor, Catho, LinkedIn), cuyos requisitos incluyen análisis de datos, uso de taxonomías u otros sistemas de organización del conocimiento (SOC) y formación en Ciencias de la Información. Las competencias extraídas se procesaron con el apoyo de OpenAI y se organizaron en ocho macrogrupos: gestión empresarial, gestión de datos, organización de la información, análisis estadístico, programación, tecnologías web, capacidad analítica y temas específicos. Los resultados indican que las empresas valoran las habilidades técnicas e interpersonales de los IP, además de su familiaridad con estándares, software y representaciones semánticas. Sin embargo, también exigen una sólida formación en estadística y liderazgo. Se concluye que los IP desempeñan un papel relevante y competitivo en el mercado de datos, lo que exige una actualización profesional constante.

Palabras clave: ciencias de la información; profesional de la información; mercado laboral; ofertas de empleo web; habilidades; experiencia; análisis de contenido; openai; análisis de datos; organización taxonómica.

1 INTRODUÇÃO

Declaração de conceito e de aplicação. Se parte de que os dados são decisivos para projetar as estratégias a serem seguidas na gestão empresarial e institucional e para apoiar o desenvolvimento dos projetos, em particular os de pesquisa. Ao mesmo tempo em que a transformação digital está mudando os perfis de atividade dos profissionais das bibliotecas e da informação, em especial para se integrar a equipes interdisciplinares de análise, em concreto de dados. Essa atividade é apoiada por sua competência técnica para mediar com os usuários, bem como por sua familiaridade com as TICs para processar dados científicos, governamentais ou comerciais, fazer uso de SOC formalizados e interagir com usuários internos e externos.

A abordagem hipotética decorre do reagrupamento epistemológico do conhecimento, no qual se baseia a natureza interdisciplinar da informação e, portanto, também de suas aplicações. Isso nos leva a indagar se as técnicas de CI têm continuidade na análise de dados e como elas são usadas em espaços interdisciplinares que exigem competências interdomínios e uma compreensão profunda do campo de aplicação. Se a Ciência da Informação está passando por um momento epistemológico de reunificação do conhecimento, seus profissionais são obrigados a colaborar na compreensão de fatos complexos.

A interdisciplinaridade é o referencial conceitual e de trabalho e, portanto, também de formação universitária. Portanto, esta proposta faz sentido para descobrir as razões pelas quais as empresas e organizações confiam no bom desempenho dos profissionais de biblioteconomia e informação quando se trata de analisar e organizar dados. Isso implica demandas específicas das empresas e organizações detentoras de dados sobre esses profissionais para incluí-los entre os fatores decisivos em seu relacionamento com clientes e usuários. Em suma, implica um monitoramento rigoroso do valor que os bibliotecários e profissionais da informação agregam à análise e a gestão dos dados. Por outro lado, significa se concentrar na base da segurança que os leva a participar de licitações competitivas contra concorrentes treinados ou especialistas em ciências básicas. Para contrastar essas afirmações hipotéticas, é necessário observar o que o mercado de trabalho de LIS oferece em termos de análise e organização taxonômica de dados.

A hipótese se deriva de trabalhos anteriores que observaram como as empresas e organizações estão contratando formados em Ciência da Informação para análise e gestão de dados em equipes interdisciplinares. O objetivo é caracterizar as competências, a experiência e o treinamento que os empregadores exigem dos profissionais de CI nas ofertas de emprego anunciadas na Web. Além disso, descobrir por que confiam nesses profissionais como analistas e gestores de dados.

Metodologicamente é selecionado o conteúdo de 84 anúncios de portais internacionais de emprego (ALAJobLIST, Glassdoor, Catho, LinkedIn) que contém tanto os conceitos de análise de dados quanto de taxonomia ou outro SOC. E, como treinamento explícito, qualquer um dos cursos de Ciência da Informação. Os anúncios coletados são marcados com XML para delimitar a descrição, as competências e os requisitos. As competências se processam então com OpenAI para extrair o conjunto de suas palavras-chave. Sua heterogeneidade torna necessário padronizá-las definindo manualmente as equivalências. A lista de sinônimos resultante é identificada e classificada por clusters temáticos do OpenAI. Propõe igualmente macroclusters para categorizar os clusters.

Por fim, são realizadas as operações necessárias para obter múltiplas estatísticas. Em primeiro lugar, através de tabelas, são descritas as circunstâncias de localização, formação, nome das vagas e identificação das entidades que publicam os anúncios. Em seguida, as competências e habilidades requeridas são comentadas e discutidas a partir dos clusters de sinônimos. Para dar congruência à análise, estes são sistematizados em oito macroclusters: Gestão empresarial e de negócios; Gestão de dados; Gestão de informações, biblioteconomia e organização do conhecimento; Análise de dados e estatísticas; Programação e automação de escritório; Padrões e tecnologias da Web; Habilidades analíticas e criatividade; e Temas de domínios específicos.

A pesquisa apresenta como principais conclusões que o Profissional da Informação trabalha em equipe com outros especialistas para analisar dados e organizar resultados de SOC semânticos. Isso porque as competências técnicas, tecnológicas, em especial as relacionadas a softwares, e de navegação estão bem contempladas nos componentes curriculares e são confirmadas na prática profissional. Dada a natureza empresarial e comercial dos empregos anunciados, é sublinhada a necessidade de competências de gestão, operacionais e de liderança, bem como de habilidades interpessoais, de comunicação e de expressão, nomeadamente na relação com os clientes. A formação e o desempenho dos PI na análise de dados merecem a confiança das empresas multinacionais pelo domínio de padrões e métodos de descrição, representação, visualização e transmissão, mas não é suficiente. Se necessita formação sólida em estatística, que já é fornecida por muitos currículos atuais. Caso contrário, é necessária uma formação no local de trabalho ou uma grande destreza no manejo das interfaces das plataformas de análise.

2 CENÁRIO

Os dados geram conhecimento e seu tratamento segue uma racionalidade clara cuja análise é usada para tomar as decisões corretas. Essas premissas tornam os dados decisivos para projetar adequadamente as estratégias a serem seguidas na gestão empresarial e institucional e para apoiar o desenvolvimento de todos os tipos de projetos, especialmente os de pesquisa. O estudo dos dados se realiza geralmente por pessoal das áreas integradas STEM, que analisam os dados em qualquer organização ou empresa para descobrir padrões e tendências e identificar situações atípicas (Štuikys; Burbaitė, 2024).

A Ciência da Informação se arquitetou em uma época de mudanças comunicativas e tecnológicas, moldada por várias contribuições disciplinares (Bereijo, 2019). Desde então o computador permite acesso aos dados, registros processados dos resultados das pesquisas, amostras e provas com particular atenção às listas de nomes e datas, códigos e cálculos que se tornaram outro suporte dos sistemas ao codificar e expressar informações de forma quantitativa, objetiva e precisa (Crisci; Armengol, 1983). O interesse dos profissionais da informação em dados os levou, de início, a assumir a gestão de projetos de pesquisa (Semeler; Pinto; Rozados, 2019). Logo, agir com Big Data significou abordar as habilidades necessárias na Ciência de dados e como adquiri-las, para continuar gerenciando as necessidades de informação dos usuários. Mas não podiam limitar-se aos dados científicos, pois os serviços analíticos se estendem ao valor informativo que revela tendências, desde as que se planificam e tomam decisões todo tipo de administrações públicas, empresas, indústrias ou entidades sem fins lucrativos (Herndon, 2021). A eficiência das organizações aumenta com a gestão apropriada dos dados, que devem ser processados de maneira contínua para não perder a capacidade preditiva e permitir que surja o conhecimento.

Por outro lado, a transformação digital está mudando os perfis de atividade dos profissionais da informação. Nessa conjuntura, aumentam as ofertas de emprego para integrar as equipes de colaboradores de ambientes diversificados que se singularizam em serviços de análise de dados, pois exigem, entre outros diplomas universitários, os de LIS. A oportunidade está condicionada a ter as competências necessárias para trabalhar como analistas de dados ou bibliotecários, taxonomistas, arquitetos de informações e gerentes de conhecimento, conteúdo ou inteligência competitiva (Paletta; Moreira, 2021). Mesmo que as ofertas não reduzam a qualificação exigida ao histórico acadêmico, pois ela também é derivada das habilidades e da experiência adquiridas em relação à preparação, à análise e a gestão da qualidade e do uso dos dados. Sem dúvida, esse novo perfil apresenta uma clara condição informativa, pois, além de

identificar, selecionar e classificar os dados, garante sua rastreabilidade, salvaguarda e acesso, com base na experiência na gestão de hierarquias e de objetos de conteúdo e na aplicação normativa de metadados que descrevem os dados e garantem a eficácia de uso. Por isso é essencial ter uma perspectiva para que a coleta, integração e processamento preservem seu contexto. Porém, muitas organizações não possuem metadados ou, quando os tem, são deficientes ou não-estruturados. Com frequência os dados se espalham por toda a organização, com informações idênticas aparecendo desagregadas e até geridas com distintos vocabulários (Agarwal; Dhar, 2014).

A identificação da ideia de registrar o crescimento dos profissionais de biblioteca e informação envolvidos na gestão e acesso a dados deve-se a Stuart (2020). Embora focado nos desafios enfrentados pela Ciência e pelos cientistas de dados. No entanto, ele destaca como as bibliotecas estão cada vez mais envolvidas na prestação de serviços entre as pessoas e o cenário de dados gradualmente complexo. Isso requer as habilidades certas, agora reforçadas pelo desenvolvimento de softwares e ferramentas de código aberto que facilitam cada vez mais a escrita e o compartilhamento de códigos. Até prevê o impulso que a análise de *big data* e a inteligência artificial trazem para esse fim. No entanto, reconhece que seu livro aborda apenas as possibilidades de emprego oferecidas pelas crescentes quantidades de dados (Stuart, 2020).

O cerne desses perfis recentes está na análise e no cuidado com os dados: sua linhagem, dados mestres, qualidade, ética, governança, arquitetura e modelagem, juntamente com a classificação granular e o processamento em taxonomias (Burton *et al.*, 2018). Sempre a partir de um design que ofereça suporte à acessibilidade e à navegação dos dados e produtos visualizados, além de proporcionar experiências envolventes e significativas para os usuários. Se opera com métodos estatísticos e algoritmos para agrupar e com aprendizagem automática para classificar. O agrupamento se efetua desde coeficientes de similaridade, mais outros de diferença, correlação e associação (Stuart, 2020). Já para a modelagem e tomada de decisões se usam métodos preditivos e algoritmos genéticos (Griffey, 2019). Apesar de que os dados geridos não sejam do volume dos *Big Data*, cuja análise se faz no convívio de recursos estatísticos de inteligência artificial como *machine learning* (ML), *deep learning* e análise de textos (Semeler *et al.*, 2024). É oportuno salientar a intervenção da IA na automação e a recuperação pois, além de classificar com padrões de aprendizagem, opera entre os usuários e os conjuntos de dados cuja disponibilidade, qualidade e acessibilidade aprimora, assim como o seu processamento e reutilização.

Para executar essas diferentes tarefas, há muitas capacidades, competências e habilidades que devem ser dominadas para que essa intervenção profissional seja confiável. E, conseqüentemente, também para o treinamento que a habilita. É claro que essa ocupação dos profissionais das bibliotecas e da informação não tem pés de barro. Pois é respaldada por um aperfeiçoamento técnico que os torna ideais para construir pontes entre os usuários e os dados científicos, governamentais ou comerciais. De uma perspectiva empresarial e institucional se consideram próprias dos profissionais da informação as capacidades de analisar e avaliar os dados (Hair *et al.* 2009). Como resposta, os bibliotecários de dados e os taxonomistas digitais gerenciam e organizam o conhecimento a partir da teoria de organização, recuperação e sistemas de informação. Isso permite que eles trabalhem em vários ambientes que exigem habilidades em metadados e gestão de dados e de objetos de conteúdo. Os esquemas de metadados e os SOC contribuem para o controle, acessibilidade, interoperabilidade e, portanto, análise e exploração de dados. Organizar os resultados da análise de dados envolve ir além das técnicas de classificação estatística, estendendo-se ao uso de SOC para gerenciar a governança, o acesso e o uso de dados e produtos. Para isso, contam com as ferramentas oferecidas pelas TICs para criar e manter taxonomias na Web, seja para intranets de empresas ou para relacionamentos com usuários e clientes ou com outras organizações por meio da Internet. Os SOC refletem os objetivos da organização intelectual de um campo de conhecimento ou de uma instituição, o que os leva a serem diferenciados de acordo com o tipo de entidade ou empresa e os dados a serem processados, embora estejam sempre muito próximas do mundo real cuja semântica representam e da perspectiva dos usuários a quem servem.

Em pesquisas anteriores, se tem atendido o mercado de trabalho brasileiro de Ciência da Informação afetado pela transformação digital, tanto no setor público quanto no privado. Agora o ciclo se fecha com o estudo amplamente internacional de vagas abertas não apenas para graduados ou profissionais de CI, muito menos para um único perfil de arquivista, bibliotecário ou gestor da informação. A persuasão para fazer isso decorre da abordagem epistemológica de reagrupamento do conhecimento, cujo princípio mais estável é que, se a natureza da informação é interdisciplinar, suas aplicações também o são (Saracevic, 1995). As profissões da informação se situam diretamente na interseção das competências e dos aplicativos de TIC e, portanto, refletem as conseqüências inevitáveis da transformação digital que renovou profundamente os paradigmas e as práticas profissionais. Nesse sentido, é necessário investigar se as técnicas específicas de CI têm continuidade e como elas são usadas em condições determinadas pela necessidade de competências comuns a diferentes setores de

atividade. Isso porque o trabalho é realizado em equipes e espaços interdisciplinares em tarefas integradas que exigem competências situadas em uma encruzilhada de domínios (Moreiro, 2017).

Como decorrência, a própria gestão digital leva a uma combinação de áreas para atender à natureza transversal dos conteúdos e das competências, tanto em termos de treinamento quanto de aplicação profissional. Além disso, nesse contexto, qualquer estudo deve se adequar a trabalhos em estruturas organizacionais interdisciplinares com bons recursos tecnológicos. Conseqüentemente, essas aspirações se refletem na necessidade de analisar a capacidade de gerenciar grandes quantidades de dados, a atenção onipresente ao usuário e às finalidades institucionais, o manuseio avançado de ferramentas TICs e uma compreensão profunda do ambiente de aplicativos. É necessário entender os perfis a serem abordados em uma estratégia empresarial ou institucional em que a comunicação e a gestão de conteúdo digital são combinadas. Além disso, para saber como os profissionais da informação contribuem para a análise de dados, o mecanismo de processamento neste artigo é a interface de programação de aplicativos OpenAI, pois facilita os procedimentos de análise do conhecimento coletado tanto das contribuições básicas quanto da organização das competências identificadas. Não há melhor maneira de incluir algoritmos no processamento de dados.

Se a Ciência da Informação está passando por um momento epistemológico de reunificação do conhecimento, seus profissionais são obrigados a colaborar e ajudar na compreensão de fatos complexos, ainda mais em um contexto de transformação científica e tecnológica. A interdisciplinaridade é a referência conceitual e laboral e, portanto, também a base da educação universitária. Os conteúdos educativos precisam ser abordados com uma resposta comparável das universidades que não podem arcar com currículos desatualizados (Bolaños; Bueno; Alves, 2013).

A proposta aqui apresentada faz sentido ao buscar descobrir as razões que sustentam a confiança das empresas e organizações no bom desempenho dos profissionais de biblioteconomia e informação no que se refere à análise e à organização de dados. Isso implica demandas específicas das empresas e organizações que detêm os dados sobre esses profissionais para incluí-los entre os fatores decisivos em seu relacionamento com clientes e usuários. Em suma, implica um monitoramento rigoroso do valor que os bibliotecários e profissionais da informação trazem para a análise e o gerenciamento de dados de informação de textos e outros objetos de conteúdo. Por outro lado, significa concentrar-se na base da segurança que os leva a participar de licitações competitivas contra concorrentes treinados ou

especialistas em ciências básicas. Se quisermos dar um contraste verdadeiro a essa exposição, a confirmação vem do aprofundamento desses e de outros argumentos hipotéticos do estudo do mercado de trabalho em LIS relacionado à análise e à organização taxonômica dos dados. Esse é o contexto desta proposta, que visa a obter valor generalista a partir de princípios aplicados.

As empresas consideram próprias dos PI as capacidades de analisar e avaliar os dados (Hair *et al.* 2009). Para o que resulta essencial ter habilidades nas TIC associadas a análise e interpretação de dados (Noy, 2014). Se opera com métodos estatísticos e algoritmos para agrupar e com aprendizagem automática para classificar. O agrupamento se efetua desde coeficientes de similaridade, mais outros de diferença, correlação e associação (Stuart, 2020). Já para a modelagem e tomada de decisões se usam métodos preditivos e algoritmos genéticos (Griffey, 2019). Apesar de que os dados geridos não sejam do volume dos *Big Data*, cuja análise se faz no convívio de recursos estatísticos de inteligência artificial como aprendizado de máquina (*ML*), aprendizado profundo e análise de textos (Semeler *et al.*, 2024). É oportuno salientar a intervenção da IA na automação e a recuperação pois, além de classificar com padrões de aprendizagem, opera entre os usuários e os conjuntos de dados cuja disponibilidade, qualidade e acessibilidade aprimora, assim como o seu processamento e reutilização.

3 REPRESENTAÇÃO TAXONÔMICA DOS DADOS

A finalidade do processo analítico dos dados não é só descritiva e de diagnóstico, mais preditiva. Com assiduidade, os dados são representados, estruturados e geridos em taxonomias que harmonizam categorização com visualização (Gani *et al.*, 2015). O que as ratifica como protótipo sistematizador, mesmo que longe do caráter terminológico inicial. Deste modo a esquematização dos dados pode gerar novos conhecimentos, mas nem sempre são analisados do jeito correto para fazer uma síntese visual que facilite a interpretação.

Trata-se de descobrir padrões e tendências, de facilitar a pesquisa ou a tomada prospectiva de decisões (Husamaldin; Saeed, 2019) que requer a colaboração de gestores de dados e de metadados, junto com os próprios pesquisadores ou responsáveis. Para interpretar é essencial conhecer as características da organização na qual os dados analisados se aplicam (Cukier; Mayer; De Véricourt, 2021). Quando os dados são empresariais se entendem desde as regras de negócio. Definir bem uma taxonomia é fundamental na hora de fornecer relevância para a etiquetagem de dados e para melhorar a exploração, acesso e compreensão intuitiva pelos usuários e clientes. Seu serviço para organizações muito diferentes faz que somente conhecendo o campo específico seja possível distinguir o que é relevante, em maior grau nos *Big Data*.

O mais frequente é visualizar os dados num conjunto gráfico, sucessor da árvore taxonômica (ANSI/NISO, 2005), que reconheça o mais significativo e norteie a navegação pelos dados, relações e tendências estatísticas desde as que minimizar a incerteza e permitir a interpretação. Os grafos de conhecimento configuram a arquitetura de tratamento desde os padrões da Web Semântica que aportam caráter ontológico e concedem interoperabilidade aos metadados e dados vinculados (Bonatti *et al.*, 2019). Os desenvolvedores são intermediários estratégicos que asseguram a comunicação direta entre as unidades organizacionais e a interoperabilidade dos recursos simbólicos: objetos, conhecimentos, informações, termos ou dados, de modo a propiciar a transparência e rastreabilidade dos processos de tomada de decisão (Lomax, 2019).

3.1 CONTEXTO E OBJETIVOS

Á medida que as TIC melhoram o processamento é progressiva a convergência com a Ciência de dados, pelo que teve que ser comprovada a demanda de bibliotecários acadêmicos preparados em dados (Federer, 2018), (Khan; Du, 2018). Bem como se constatou sobre anúncios de emprego que, em apenas cinco meses, a habilitação bibliotecário-cientista de dados empresariais oferecia no Brasil 12 vagas, enquanto eram 11 as de analista-gestor de dados e 6 para analista-gestor de dados em projetos (Paletta; Moreiro, 2021). Estes valores indicam que as organizações estão integrando os profissionais da informação em equipes interdisciplinares para a gestão de dados, por suas competências na descrição, análise, organização e difusão. Assim como, ao oferecer oportunidades profissionais para formados de diferente origem, confirmar até que ponto abre o espectro de contornos para os que podem optar os titulados em Ciência da informação. Como também conhecer as características de formação, experiência e ocupação envolvidas numa capacitação heterogênea, por difícil que seja agir como analista de dados (Oliver *et al.*, 2019). Esta co-ocorrência vem favorecida pelo propósito estratégico dos SOC de caracterizar os contextos de uso e atender às necessidades reais e práticas.

O objetivo geral é verificar a penetração de perfis caraterísticos da análise de dados nas habilidades profissionais da CI. Como aquisição cognitiva, supõe verificar os requisitos estabelecidos pelas empresas para ser analistas e sistematizadores de seus dados. Consoante com compreender como os empregadores definem um PI de dados através das qualificações que declaram e listam, nas que coexistem competências próprias da análise e da gestão de dados. Assim como perceber sua organização e gestão mediante esquemas taxonômicos e outros SOC distintivos de um campo, empresa ou entidade.

3.2 METODOLOGIA

Se utiliza a análise de conteúdo para verificar as evidências por trás da seleção de 84 anúncios a texto completo em portais de emprego internacionais. As vantagens da análise de anúncios web para a compreensão das atividades realizadas pelos PI foram destacadas num trabalho anterior (Paletta; Pastor; Moreiro, 2021). Agora, os anúncios têm que conter, em coincidência, a análise de dados, junto a taxonomia ou outro SOC e até mesmo, como formação explícita, qualquer dos cursos em Ciência da Informação. No início se escolhe *LinkedIn* como fonte não apenas por ser uma ampla rede para ofertas de emprego, mas devido à existência do grupo de discussão *Taxonomy and Ontology Community of Practice*, cujos administradores são especialistas em gestão de dados. Além disso, ainda é utilizada a *ALAJobLIST* que oferece a vantagem de direcionar os anúncios para profissionais ou formados em CI; bem como o *Glassdoor Jobs* de caráter internacional e com descrições detalhadas, e o site brasileiro *Catho*. As buscas se fazem no último mês de 2022 e nos três primeiros de 2023 através dos termos Análise de dados e taxonomista, Análise de dados e taxonomia, Taxonomia de dados, Arquiteto de metadados e Análise e organização de dados.

Tabela 1 - Fontes dos anúncios

| Rede profissional | No de ans | % |
|--------------------------|------------------|----------|
| LinkedIn | 33 | 39,28 |
| Glassdoor Jobs | 31 | 36,90 |
| ALAJobLIST | 11 | 13,09 |
| Catho | 9 | 10,71 |
| Total | 84 | 100 |

Fonte: dados da pesquisa

São apontadas apenas as ofertas nas que coincidem análise de dados com sua organização taxonômica. No período coberto, outros 97 anúncios de LinkedIn ofertaram vagas de analistas de dados, principalmente para modelagem, que requerem formação superior em STEM (ciência, tecnologia, engenharia e matemática, de acordo com seu acrônimo em inglês), não considerando a Ciência da Informação, pelo que ficaram de fora. Na ALAJobLIST se publicaram 79 ofertas para análise e visualização de dados, porém, menos da metade estava relacionada à organização de dados usando taxonomias ou outros SOC. Como o tratamento é feito em inglês, algumas ofertas da Catho se têm traduzido. Bem como se eliminaram 34 ofertas de trabalho duplicadas. A amostra selecionada se considera suficiente para ser representativa e para obter a reprodutibilidade necessária que torne confiável seu conteúdo (Krippendorf, 2018).

A análise se fez em duas fases: uma primeira tabular, de natureza enumerativa e descritiva dos resultados e as entidades identificadas automaticamente. A segunda avalia qualitativamente o conteúdo dos anúncios, após o seu processamento. Na fase primeira se dão essas sete etapas:

Passo 1. Depois de obter os anúncios em formato de texto e no idioma inglês, os documentos são estruturados de acordo com o seguinte esquema XML para torná-los homogêneos e expressivos antes da análise (Berelson, 1952):

- <content> : Elemento raiz do documento XML.
- <content>→<main> : Anúncio individual.
- <content>→<main>→<description> : Descrição geral do anúncio, incluindo título, local de publicação e texto informativo.
- <content>→<main>→<tasks> : Tarefas a serem executadas na oferta de emprego.
- <content>→<main>→<requeriments> : Qualificações e competências relacionadas a oferta de emprego.
- <content>→<main>→<requeriments>→<qualifications> : as qualificações exigidas se especificam em linhas diferentes.
- <content>→<main>→<requeriments>→<competences> : as competências exigidas se especificam em linhas diferentes.

Passo 2. A API OpenAI é usada para extrair as palavras-chave originais das competências requeridas em cada oferta de emprego. O modelo text-davinci-003 é aplicado com o seguinte prompt:

Extract keywords from the following list of competencies in a job offer and display them separated by semicolons.

Depois de processar cada oferta de emprego com este prompt, obtém-se um ficheiro de texto no qual as 1.554 palavras-chave extraídas das competências de todos os anúncios aparecem em linhas diferentes e separadas por ponto e vírgula. Após evitar as diferentes repetições foram obtidas 1.126 palavras-chave únicas.

Passo 3. As palavras-chave são processadas com um script Python para obter as seguintes estatísticas:

- Nome das palavras-chave
- Frequência total de ocorrências (a mesma palavra-chave pode aparecer várias vezes no mesmo anúncio). - Número de anúncios diferentes em que o termo aparece.

O ficheiro CSV resultante é revisto a mão para agrupar os sinónimos. Como exemplo, “*categorization and classification*” foi usado como sinónimo para estas palavras-chave:

Categorization; categorization; Categorization design; classification; Classification; Classification definitions; Classification schemes; Classification Systems; Classifying.

Se descartam 102 palavras-chave que não faziam sentido. Dessa forma, o número de palavras-chave selecionadas é de 1.024. As palavras-chave são levadas então a sua forma canônica (termos compostos, singular e plural) e se definem os termos preferidos (*synonyms*) correspondentes entre as palavras com significado igual ou muito semelhante. O número total de sinônimos é de 480, o que representa uma média de 2,13 palavras-chave cada termo preferido.

Passo 4. Ainda, a API OpenAI se usa para identificar *Clusters* a partir do agrupamento das palavras-chave obtidas na etapa 2. Para o efeito, se aplica o prompt:

Identify up to a maximum of 20 thematic clusters from the following list of keywords.

Quando o OpenAI retornou o número de *Clusters*, pediu-se que atribuísse cada uma das palavras-chave a um ou mais deles (até um máximo de três). O número de *clusters* propostos pela OpenAI foi de 36, que se decidiu limitar a 20 após uma análise cuidadosa e por razões de eficiência do tempo de processamento. Por que não foi pedido que fizesse isso com os sinônimos? Porque se queria manter o significado semântico original dos anúncios e porque o mapeamento entre as palavras-chave e os sinônimos já estava feito, então era automático elaborar o *clustering* de sinônimos a partir do *clustering* de palavras-chave.

Passo 5. Depois de obter a lista de grupos temáticos, a API de OpenAI se utiliza novamente para agrupar cada uma das palavras-chave do passo 2 com cada um dos grupos temáticos obtidos no passo 4. Para proporcionar alguma flexibilidade, cada palavra-chave pôde ser atribuída a dois grupos de assunto, utilizando o *prompt*:

For the following keyword: {Key_word}

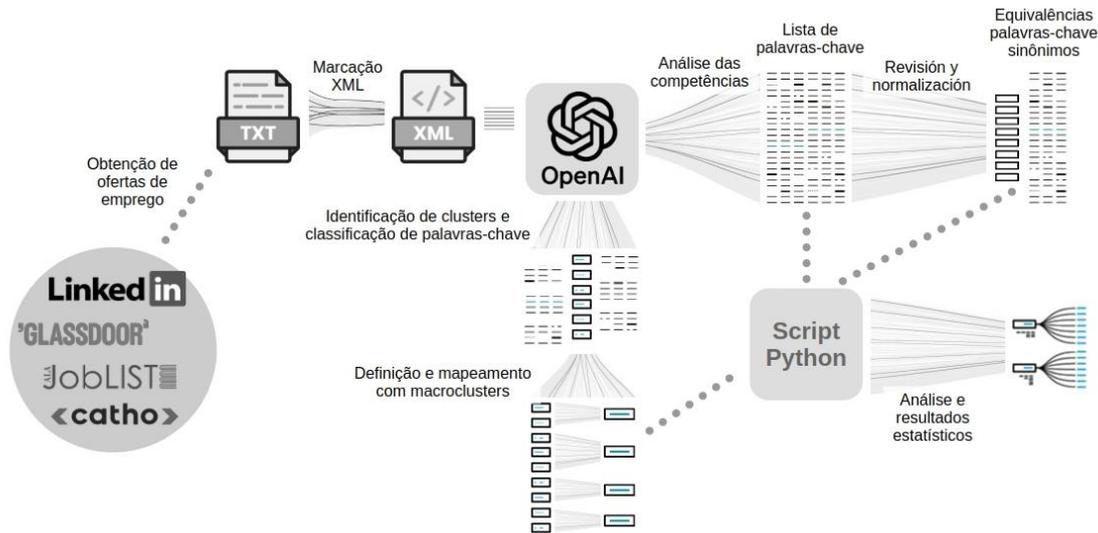
assign one of these categories (maximum 2 categories): {Subject_categories_list}.

Show only the assigned categories separated by ';'. Don't include any explanation.

Portanto, o processamento é feito caso a caso por palavra-chave. A utilização neste passo das palavras-chave originais, em vez dos seus sinônimos, permite uma maior flexibilidade se, a dada altura, for considerado desejável definir sinônimos diferentes. Por outro lado, o passo 5 se realiza utilizando os grupos temáticos do passo 4, em vez de trabalhar

diretamente com áreas temáticas mais amplas. Esta última opção facilita sempre a definição de diversas áreas para combinar os diferentes grupos.

Figura 1 - Fluxo de processamento dos dados



Fonte: Elaboração própria.

Passo 6. Se unificam os clusters a mão em 8 macroclusters. Por exemplo, o campo Gestão de dados unifica em coerência e coesão os clusters temáticos: Gestão de dados; Bases de dados; Gestão de projetos; Segurança e privacidade e Governança e normas.

Passo 7. Por fim, se utiliza um script Python para calcular os dados estatísticos agregados sobre a frequência de anúncios por grupo temático e área.

Passo 8. A interpretação dos dados obtidos se foca no comportamento das ofertas de emprego no campo da análise de dados para PI.

4 RESULTADOS DE CARÁTER DESCRITIVO

Nesta primeira fase se identificam as características gerais da amostra estudada por meio de tabelas comentadas, bem como se observa o comportamento das circunstâncias de lugar, formação, posição e corporação.

Tabela 2 - Países ou regiões onde a atividade profissional é realizada.

| País | Nº de vagas de emprego |
|----------------|------------------------|
| EUA | 27 |
| Inglaterra | 14 |
| Brasil | 12 |
| Canadá | 5 |
| Espanha | 4 |
| América Latina | 3 |
| Colômbia | 2 |
| França | 2 |
| Norte América | 2 |
| Reino Unido | 2 |
| Outros | 11 |
| Total | 84 |

Fonte: dados da pesquisa

Outros incluem 11 países com uma única ocorrência: Argentina, Índia, Irlanda, Cazaquistão, México, Polônia, Cingapura, Coreia do Sul, Suécia, Suíça e Turquia. Percebe-se um viés interno no caso do Brasil, pois foram incluídos os 9 anúncios que apareceram na rede Catho, além de 3 de outras fontes. Seis anúncios de emprego (7,14%) são oferecidos para trabalho remoto, três dos quais nos EUA, dois na América Latina, postados dos EUA, e um na Colômbia.

15

Tabela 3 - Nomes e números de títulos universitários em LIS

| BCI | | MCI | |
|-----------------------------------------|----|-----------------------------------------|----|
| Arquivos | 2 | Gestão da informação digital | 4 |
| Gestão da informação | 4 | Organização do conhecimento | 2 |
| Ciência da informação | 8 | Gestão da informação | 3 |
| Estudos da informação | 4 | Ciência da informação | 9 |
| Biblioteconomia e ciência da informação | 10 | Biblioteconomia e ciência da informação | 8 |
| Biblioteconomia | 12 | Dados da biblioteca | 2 |
| Gestão de documentos | 4 | Biblioteconomia | 8 |
| | | Serviços de biblioteca | 2 |
| | | Metadados | 2 |
| | | Taxonomia/Ontologia | 4 |
| DCI | | Webmetria | 2 |
| Ciência da Informação | 1 | | |
| Total | 45 | | 43 |

Fonte: dados da pesquisa

Os anúncios apresentam uma ampla variedade de designações para os cursos universitários de LIS que a tabela 3 mostra como bacharelado (BCI); mestrado (MCI) e doutorado (DCI). Entretanto, o currículo não é mais apenas o diploma acadêmico, mas o

acumulado de habilidades e experiências adquiridas. Por isso, além do título, 35 das ofertas exigem anos de experiência, 3 anos em média para bacharelados e 4,5 anos para mestrados. Até quatro deles validam o MCI para um número equivalente de anos de experiência após a graduação. Dois anúncios exigem que os estudos tenham sido feitos em uma faculdade ou universidade credenciada. Em outros dois, um mestrado credenciado pela ALA. Outros quatro permitem uma formação de tempo equivalente em informação, ciência de dados ou taxonomia, como uma combinação equivalente de formação e experiência. Seis vezes é fixado que “preferencialmente mestrado”. Ocasionalmente, o diferencial não está nos estudos realizados, mas nas competências estatísticas (6), de programação (2) ou de domínio de software específico (5).

Tabela 4 - Organizações comunitárias e unidades de informação que originam posições

| Organizações comunitárias | Unidades de informação |
|---------------------------------------------|------------------------------------------------------------|
| British Horse Society | Bibliotecas IBRM (São Paulo-BR) |
| Epilepsy Foundation. MD-USA | McLeod Business Library. VA-USA |
| Lausanne Switzerland Information Technology | The Felipe Herrera Library. IADB |
| Metropolitan Council. Saint Paul, MN-USA | The National Archives, England (2 Ofertas de trabalho) |
| Observatorio de Drogas de Colombia | Raymond A. Mason School of Business. VA-USA |
| | SOAS (Library Records) University of London |
| | Stockholm Environment Institute |
| | University of London |
| | University of Nevada |
| | Wilfrid Laurier University. ON-CAN (2 Ofertas de trabalho) |
| Total: 5 | Total: 12 |

Fonte: dados da pesquisa

As ofertas para centros de informação registram o impacto das bibliotecas e dos arquivos desde uma perspectiva empreendedora de seus serviços, objetivos e desempenho. Os arquivos e as bibliotecas universitárias detalhadas gerenciam, vinculam e preservam ativos digitais pelo processamento dos dados de seus recursos, especialmente de projetos e atividades de pesquisa, além de avaliar o desempenho da instituição e facilitar a tomada de decisões informadas. Graças ao fornecimento de serviços a partir de objetos digitais com metadados vinculados para

compartilhar sua experiência externamente e dialogar com o feedback dos usuários. Embora seja certo que exigem perfis enquadrados no campo LIS em 23 das ofertas (26,1%).

Tabela 5 - Empresas que originam as ofertas de trabalho

| Principais empresas contratantes | Empresas com mais de uma posição | |
|-----------------------------------------|-----------------------------------------|--------------|
| | <i>Company</i> | <i>Vagas</i> |
| 3CON | HP | 5 |
| Airbus | BairesDev | 4 |
| Amazon | Bloomberg | 3 |
| AstraZeneca | Expedia | 3 |
| Boeing | JPMorgan Chase Bank | 3 |
| Compass UOL | Schneider Electric | 3 |
| Deliveroo | Curated Real experts | 2 |
| Deloitte | Netflix | 2 |
| Delta | | |
| Dow Jones | | |
| HireTalent | | |
| 3CON | | |
| IDB Inter-American Development Bank | | |
| INL | | |
| MediaMonks | | |
| Meta | | |
| Petrobras | | |
| PwC | | |
| Resideo | | |
| Standard & Poor's Global 100 | | |
| UBS | | |

Fonte: dados da pesquisa

Nos 79,7% dos anúncios que identificam as empresas contratantes, destacam-se as grandes empresas multinacionais. Que sejam contratados por empresas tão significativas ressalta a confiança na sua capacidade para, além de administrar os ativos digitais, desenvolver os processos relacionados com a análise, acessibilidade e qualidade dos dados e sua organização taxonômica. Trabalham em qualquer empresa e setor de atividade. 4% dos anúncios são feitos para “Empresa confidencial”.

Como detalhes a serem considerados nessa enumeração, as três ofertas da *Schneider Electric* são idênticas, embora tenham sido feitas separadamente para atividades na França, Polônia e Espanha. A empresa de software BairesDev está localizada na Virgínia (EUA), embora seus anúncios de emprego tenham sido dois para a América Latina, um para o Brasil e outro para a Colômbia.

Tabela 6 - Nomes de cargos profissionais em centros de informação

| Arquivos | Bibliotecas |
|--------------------------------------------------|----------------------------------------------|
| Assistente de Arquivos | Analista de dados de biblioteca |
| Especialista em Informação e Recursos | Bibliotecário de análise de dados e negócios |
| Gestor de Mídia Digital (2) | Bibliotecário (2) |
| Indexador de Conteúdo | Bibliotecário associado |
| Especialista em Ativos Digitais | Bibliotecário de ativos digitais |
| Arquivista de Projetos | Bibliotecário de dados |
| Especialista em Gestão de Informação e Registros | Bibliotecário de recursos |
| Especialista em gestão de Documentos | Bibliotecário taxonomista e de metadados |
| Analista de Informação/Gestão de Dados (2) | Bibliotecário técnico |
| Especialista em Geração de Informação | Chefe de catalogação, taxonomia e dados |
| Gestor de Informação | |

Fonte: dados da pesquisa

Quais são as ocupações dos PI? Embora mantenham a designação principal original, até mesmo seis desses nomes são dedicados ao processamento, organização e visualização de dados empresariais e comerciais. Ainda assim, quase 80% das denominações de vagas restantes não utilizam os termos tradicionais que denominam aos PI, o que reflete diretamente as alterações nos serviços prestados. Em particular na encruzilhada das empresas o comércio e as TIC se preferem outros termos no título para desenvolver umas atividades combinadas dentro de processos abertos. Perfis transversais para equipes transdisciplinares onde titulados de diferentes campos de conhecimento a partir de bons meios tecnológicos participam para criar, tratar, organizar e gerenciar dados e informações digitais.

Tabela 7 - Nomes de cargos profissionais em centros de informação

| Analista de dados | Organizador de dados | Arquiteto de informação |
|-------------------------------------------------|-------------------------------------------|---------------------------------------------------|
| Analista de proteção e privacidade de dados | Consultor de ontologia | Arquiteto de conteúdo UX eCommercio (4) |
| Gerente de conhecimento (2) | Dados de taxonomia de produto | Arquiteto de dados e informações |
| Consultor técnico de finanças sustentáveis | Gestor de produto | Arquiteto de informações |
| Analista de negócios | Associado de prototipagem de vocabulário | Arquiteto de informações digitais/taxonomista (2) |
| Analista de analíticas Web (2) | Engenheiro de conhecimento/ontólogo (3) | Coordenador de conteúdo |
| Especialista/taxonomista em gestão de dados | Consultor de ontologia/taxonomia | Designer de conteúdo UX (2) |
| Análise de dados (forense) | Especialista em dados de taxonomia ESG | Especialista em entrada de conteúdo de mídia |
| Analista de conhecimento | Estrategista de taxonomia | Estagiário de arquitetura de dados |
| Analista de texto | Coordenador de operações/projeto | Estrategista de conteúdo (3) |
| Analista de dados de transformação de taxonomia | Linguista/taxonomista de engenharia | Estrategista de conteúdo UX (2) |
| Especialista em curadoria de dados | Modelador semântico sênior | |
| Analista de negócios de TI | Ontólogo | |
| Analista de dados (4) | Taxonomista (6) | |
| Analista de dados e taxonomista | Taxonomista de dados de produto | |
| Analista de governança de dados (2) | Taxonomista de gestão de dados de produto | |
| Analista de otimização | Taxonomista global de produtos | |
| Gerente de dados e inovação | | |
| Investigador de etiquetagem de dados | | |

Fonte: dados da pesquisa

Como já observado nas designações de profissionais da informação em anúncios de emprego relacionados à transformação digital (Paletta; Pastor; Moreiro, 2021), confirma-se agora que *o analista* inclui em 18 aparições uma expressão muito variada. Não apenas o compromisso óbvio com a análise de dados, mas também com produtos, textos ou Web.

Associando-se estreitamente aos 6 rótulos ocupacionais de analista ou gerente do conhecimento. Sem dúvida, a seleção, análise e estudo de objetos de conteúdo da Web é uma das ocupações mais abrangentes e crescentes dos profissionais da informação (Agarwal; Dhar, 2014).

Enquanto organizador de dados arrola as denominações de quem administra e visualiza o conhecimento do campo de dados e produtos (Nickerson; Varshney; Muntermann, 2013), ao tempo que contempla ferramentas para a comunicação com outros analistas, mas em especial com os usuários e clientes. Em propósito completado pela categoria arquiteto da informação com as designações dos analistas de experiência de usuário (UX), assim como as de arquiteto de informação e dados, coordenador de conteúdo, designer ou estrategista dos objetos de conteúdo a serviço do comércio eletrônico.

A atuação coincidente em assuntos e tarefas misturadas requer competências e habilidades tocantes a diferentes campos e formações originárias, assim como a variados espaços de aplicação (Kennan *et al.*, 2006). Num cenário laboral de transformação digital resulta determinante a caracterização dimanante da conexão às redes de dados e o compartilhamento em tempo real do conhecimento produzido. De modo que a conveniente atuação profissional não prove só do expediente acadêmico possuído (Becker *et al.*, 2017), mas do conjunto de destrezas e experiências alcançadas.

4.1 ANÁLISE DAS COMPETÊNCIAS E CAPACIDADES REQUERIDAS

A análise avaliativa baseia-se nos 8 macroclusters inferidos, que unificam os clusters de sinônimos e, portanto, de palavras-chave enunciadas pelo Open AI. Ocasionalmente, as classificações de conceitos resultantes dessa aplicação podem parecer estranhas. Mas as propostas são respeitadas porque, mesmo com alguma subjetividade, podem ser compreendidas no âmbito do conjunto sugerido.

Tabela 8 - Estatísticas descritivas dos macroclusters.

| Macrocluster | F KW | Nº KW | Nº sinônimos | J | F |
|---------------------------------------------------------------------|------|-------|--------------|---|----|
| Gestão de negócios e empresas | 413 | 340 | 113 | | 73 |
| Gestão de dados | 226 | 181 | 88 | | 73 |
| Gestão da informação, biblioteconomia e organização do conhecimento | 228 | 151 | 68 | | 68 |
| Análise de dados e estatísticas | 198 | 139 | 78 | | 62 |
| Programação e automação de escritórios | 193 | 98 | 62 | | 60 |
| Padrões e tecnologias da Web | 194 | 116 | 68 | | 55 |
| Habilidades analíticas e criatividade | 116 | 76 | 24 | | 49 |
| Temas de domínios específicos | 47 | 48 | 42 | | 23 |

Fonte: dados da pesquisa

Os macroclusters se ordenam pela frequência nas ofertas de trabalho. A frequência das palavras-chave (KW) reflete a importância relativa de cada grupo com respeito ao objeto de estudo e ao desenho da estratégia de busca. Em termos de frequência e número das palavras-chave que representam, os macroclusters se mostram bastante equilibrados. Entretanto, no topo, destaca-se Gestão de negócios e empresas, que quase dobra o próximo e contém 25,57% da frequência das palavras-chave e 29,59% de seu número. Isso mostra o caráter empresarial e comercial da maioria das ofertas. Na parte inferior, temas de domínios específicos tem, em ambos os casos com 2,9% e 4,1%, menos da metade da frequência de aquele. A partir de então, se acercam no número de sinônimos, mesmo que certas diferenças se mantenham nos extremos. Aproximação que se incrementa na frequência com que os macroclusters aparecem nos anúncios.

Nº de Sinônimos é a frequência total ou número de vezes que um sinônimo ou termo preferido aparece nas ofertas de trabalho. Enquanto a frequência simples de sinônimos (F) é o número de vezes que os sinônimos de um macrocluster aparecem nos anúncios. Coincide com a frequência de aparição em diferentes ofertas de emprego. Até mesmo aparecem várias vezes no mesmo anúncio. 35 dos sinônimos pertencem simultaneamente (polihierarquia) a dois macroclusters da tabela 8.

4.1.1 Macrocluster de competências Gestão de negócios e empresas

Este macrocluster alcança alto desempenho em anúncios que demandam pessoas capazes de gerenciar e controlar as atividades e os recursos de empresas e organizações. Isso significa ter as ferramentas, comportamentos e competências aplicáveis às operações e estratégia de negócio, ao marketing e, é claro, às equipes e local de trabalho. O macrocluster está incluído em 73 ofertas de trabalho diferentes (86,9%). Ele foi classificado em 5 grupos intermediários de sinônimos (clusters).

Tabela 9 - Grupos de sinônimos do macrocluster Gestão de negócios e empresas.

| Cluster | Ofertas |
|-------------------------------------------------------------------|----------------|
| Gestão de negócios e empresas | 62 |
| Habilidades organizacionais, no local de trabalho e interpessoais | 62 |
| Habilidades linguísticas | 14 |
| Habilidades de liderança e gestão | 8 |
| Gestão da educação e da empresa | 1 |

Fonte: dados da pesquisa

As capacidades usadas nos processos diários são essenciais para o sucesso dos negócios e são integradas ao cluster Gestão de negócios e empresas. Os anúncios buscam os melhores candidatos, com habilidades essenciais para entender o negócio e seus limites. Combinar a análise de negócios de ferramentas e infraestrutura de dados com as melhores práticas para alcançar o desempenho dos processos operacionais de negócios, juntamente com a melhoria dos procedimentos e estratégias técnicas. Enquanto as capacidades em Business intelligence, modelagem e gerenciamento de projetos e pesquisa operacional facilitam mudanças na organização em um contexto global e híbrido. Algumas capacidades técnicas mais específicas também são reveladas, como gestão financeira, consultoria, gestão de contratos e licitações e contabilidade. A aptidão gerencial atinge o maior destaque entre os profissionais da informação em Choi e Rasmussen (2009), também em Khan, Masrek e Nadzar (2015) ao lado da liderança. Isso sem falar na importância dada aos recursos que concentram estratégias de marketing e branding, além da familiaridade com e-commerce e merchandising que promovem uma boa experiência do usuário. Essas últimas qualidades também dizem respeito ao macrocluster Padrões e tecnologias da Web.

Em ofertas de trabalho com fortes relações interpessoais há uma grande demanda por habilidades organizacionais e comunicativas, capacidade de trabalhar em equipe e habilidades de liderança. Assim, a busca por candidatos com habilidades organizacionais e processuais não é um grupo amplo, mas essencial. Grandes volumes de trabalho são atendidos com uma alta capacidade de usar o tempo, priorizar atividades urgentes e otimizar processos. Para garantir uma interação humana eficaz e produtiva, Erozkán (2013) e Zhang e Chen (2023) enfatizaram o papel da comunicação e das relações interpessoais, confirmado pelos resultados desta análise. As habilidades de comunicação são destacadas, permitindo que os participantes compartilhem, cooperem, tenham empatia e estabeleçam relacionamentos corretos nas relações pessoais. Portanto, as habilidades organizacionais, no local de trabalho e interpessoais estão intimamente relacionadas às habilidades sociais das habilidades analíticas.

Os empregadores estão procurando profissionais que possam se comunicar verbalmente e por escrito para marketing e publicidade, em reportagens externas, em notícias econômicas e até mesmo em comunicações acadêmicas e técnicas. Com essas capacidades, os profissionais ganham autoconfiança no relacionamento efetivo com usuários, consumidores e fornecedores, mas também dentro da equipe e da organização e com os colaboradores. Goulding *et al.* (1999); Erozkán (2013); Harper (2013) os classifica entre os mais exigidos em anúncios de emprego. Atenção especial deve ser dada ao cluster Habilidades linguísticas, devido aos anúncios que exigem proficiência na língua inglesa. Outros, operando em um contexto internacional com o inglês como língua veicular, nem mesmo consideram isso. A proficiência em francês e espanhol é excepcionalmente exigida, apenas uma vez em cada idioma.

A conquista de boas habilidades de relacionamento interpessoal envolve habilidades de outras pessoas, como cultura, imparcialidade, humildade e maturidade. Compartilhar experiências incentiva, às quais atitude positiva, questões delicadas ou tato podem ser adicionados para alcançar a eficácia ao se comunicar com clientes e fornecedores ou ajudá-los (Meganck; Smith; Guidry, 2020). São muito significativos nos termos e condições gerais da maioria das ofertas, a capacidade e convicção de promover a diversidade, igualdade e inclusão. Bem como uma abordagem integrada da cultura no desenvolvimento sustentável.

Muitas das habilidades necessárias para gerenciar dados precisam um ofício esperto. Os candidatos estão requisitados a uma formação multidisciplinar que lhes permita gerar valor social, econômico e cultural nas empresas. Além de manterem em dia suas habilidades acadêmicas iniciais, devem se atualizarem com a gestão da empresa ao ingressar e, logo depois, com a formação continuada. Ao mesmo tempo, devem dispor de recursos pedagógicos para

canalizar as atividades de formação desde metodologias baseadas na reflexão teórica e na aplicação prática que favorecem o aprendizado para serviços reais.

Por fim, as qualificações em Competências de liderança e gestão estão voltadas para a supervisão de grupos, projetos, tempo e prioridades, bem como para a responsabilidade na tomada de decisões. As ocorrências refletem sua relevância devido à importância para a gestão de negócios e empresas das habilidades identificadas para liderar equipes, construir comunidades, persuadir, inspirar, supervisionar, chegar a um consenso, equilibrar tarefas e, é claro, gerenciar a equipe, os objetos e os recursos. A gestão executiva e a gestão sênior são ocupações no nível mais alto da gestão da organização.

4.1.2 Competências do macrocluster Gestão da informação, biblioteconomia e organização do conhecimento

Este macrocluster congrega 81 sinônimos em dois clusters. A maioria está incluída no primeiro, de Gestão da informação e biblioteconomia, pois apenas 5 aparecem no cluster de Gestão da informação.

Tabela 10 - Sinônimos do Macrocluster encontrados em 3 ou mais Ofertas de trabalho

| Sinônimos | Ofertas | Sinônimos | Ofertas |
|-------------------------------|---------|--------------------------|---------|
| Taxonomias | 33 | Gestão de banco de dados | 6 |
| Metadados | 14 | Pesquisa | 6 |
| Categorização e classificação | 12 | Vocabulários controlados | 5 |
| Ontologias | 12 | Gestão de conhecimento | 5 |
| Arquitetura da informação | 10 | Bibliotecas | 4 |
| Gestão de conteúdo | 9 | Arquivos | 4 |
| Organização do conhecimento | 8 | Catálogo | 3 |
| Outlook | 8 | Gestão de documentos | 3 |
| Adobe | 6 | Edição | 3 |

Fonte: dados da pesquisa

É lógico que o sinônimo Taxonomia tenha uma alta frequência, pois é um dos elementos da estratégia de pesquisa usada para selecionar os anúncios. Os termos preferidos com maior frequência neste grupo também fazem parte de outros clusters. Assim, taxonomia, metadados e gestão de banco de dados em Gestão de dados; gestão de conteúdo, gestão de pesquisa e

ontologias em Padrões e tecnologias da Web. O que reconfirma a confluência transdisciplinar da CI na hora de analisar e organizar dados, informação e conhecimento (Bereijo, 2019).

Os requisitos das ofertas mostram que a gestão e a organização dos dados se baseiam em teorias de recuperação da informação. Mas os aplicativos e os ambientes de trabalho variam tanto quanto as empresas recrutadoras e os postos profissionais oferecidos que, na maioria das vezes, exigem experiência e perícia em metadados, gestão do conhecimento e dos objetos digitais. Neste caso, taxonomia é o termo preferido de um grupo de sinônimos que proporcionam semântica aos dados para entender as diferenças de significado e promover e eficácia de sua análise, interoperabilidade e reutilização. Mais também para projetar, implantar e governar os ativos institucionais por meio de uma arquitetura que modelize e valorize a informação corporativa (Federer, 2018). Ainda que em termos gerais as organizações perseguem a conexão de metadados com objetos de conteúdo, já sejam dados, pessoas, textos, produtos ou serviços, que se devem interpretar desde as necessidades dos clientes-usuários, às vezes internos. Assim como entender o modelo cognitivo das empresas que solicitam candidatos capazes de apoiar os objetivos almejados.

Por isto, os solicitantes estarão preparados para definir a arquitetura da informação na forma de mapas de sites, taxonomias vinculadas, esquemas de metadados, vocabulários estruturados, modelos de conteúdo, ontologias ou diagramas de relações de conteúdo que agreguem valor aos clientes e à organização. A gestão de dados exige capacidade para projetar, implantar e governar os principais ativos institucionais por meio de uma arquitetura que integre todos os dados corporativos e determine os atributos de serviço. Deste modo, se envolvem habilidades respeitantes à KO com arquitetura da informação, TIC e *UX Design*. Num contexto de atuação em que se precisa dominar o valor e os princípios dos modelos de dados (Given; Olson, 2003). Entre os diferentes tipos de dados as empresas notabilizam nas ofertas aqueles relacionados a produtos comerciais, cuja definição e diferenciação exatas, designação de atributos e categorização demandam abertamente.

Então, as habilidades mais demandadas servem para criar, melhorar e gerir taxonomias granulares de produtos de comércio eletrônico, em auxílio as empresas a criarem catálogos mais eficazes, também a refinar a experiência de usuário (Papadimitriou *et al.*, 2013) numa navegação certa na busca de objetos e dados, em especial a través de dispositivos móveis. Essa é a razão pela qual 34 das divulgações requerem experiência no desenho e gestão de taxonomias de dados de objetos comerciais e de pesquisa empresarial. Embora o período mais comum de prática estendida seja de três anos, em cinco das ofertas de este grupo se chegam a pedir cinco

anos e sete anos em outra. Mas não só, pois a capacidade de usar taxonomias é exigida duas vezes para organizar dados de projetos, bem como para coordenar conhecimentos bibliotecários. Outras nove oportunidades solicitam perícia na atualização ou melhora da modelagem taxonômica de dados. Complementado com 26 anúncios que, lato sensu, posicionam essa correta gestão de dados ou de *big data* na aptidão derivada do domínio de ontologias, grafos de conhecimento, esquemas de classificação e outros SOC, conforme avançaram Mayr *et al.* (2016). E vinculado nos anúncios ao exercício prévio em ferramentas e padrões semânticos (RDF, OWL, SKOS, SPARQL, ...) atendidos em outros macroclusters.

A prática prolongada é também solicitada em 24 dos anúncios concernentes a relação das taxonomias com os esquemas de metadados, em função de identificar e distinguir os dados, sua origem, atributos e estrutura contextual, assim como determinar sua integração, interoperabilidade e arquivamento. Esta capacidade de organizar dados permite aos candidatos preservarem a consistência da taxonomia com as necessidades da organização, tornando-os legíveis para os cientistas e para os usuários e consumidores. Para que os dados sejam significativos, é essencial conhecer o contexto de sua origem e uso, o que lhes dá seu verdadeiro significado. Esta é a grande contribuição das habilidades profissionais de informação para sua análise e interpretação (Seadle; Havelka, 2023). De grande afinidade com muitos dos atributos abordados no macrocluster de competências a seguir.

4.1.3 Macrocluster de competências em Gestão de dados

As habilidades de Gestão de dados se podem entender como *Master data* ou *Digital asset management* (DAM) uma vez que se aplicam a gestão e compartilhamento de objetos digitais em instituições e empresas.

Tabela 11 - Grupos de sinônimos do macrocluster Gestão de dados.

| Cluster | Ofertas |
|--------------------------------------------|---------|
| Gestão de dados | 53 |
| Gestão de projetos | 33 |
| Bases de dados | 20 |
| Governança e normas | 18 |
| Segurança e privacidade | 4 |
| Gestão de informações sobre produtos (PIM) | 3 |

Fonte: dados da pesquisa

Na maioria dos anúncios trata-se de dados relacionados às transações comerciais cujo ciclo de vida é gerenciado pelas empresas por meio da Gestão de dados de produtos. Em uma função que atende a organizar e reunir em um único sistema todos os dados e informações relacionados aos diferentes processos de desenvolvimento e fluxos de trabalho. Com suficiente qualificação e manejo de recursos que facultem para manter ao longo do tempo dados de diferentes formatos com sua qualidade, exatidão e integridade. A Integração de dados é atribuição da Governança e normas, pois a garante nos processos de extração, transformação e carga (ETL) que visam segurar a atualização e validade dos dados com relação aos fins para os que foram criados. Sempre desde a padronização de sistemas, práticas recomendadas, privacidade e conformidade de dados, e a consistência das políticas sobre as quais se faz a tomada de decisões.

Este macrocluster também envolve habilidades para a busca de padrões e o planejamento de Gestão de projetos. Em especial, para a coleta, organização, armazenamento e documentação de Gestão de dados de pesquisa, como bem destacou Kennam (2016) entre outras responsabilidades de trabalho dos bibliotecários. A gestão de projetos exige criatividade e habilidades interpessoais, mas, acima de tudo, competências gerenciais e organizacionais, quando não é comum que os PI tenham sido treinados nesses perfis, como observa Kinkus (2007). Portanto, para se tornarem qualificados, os possíveis gestores de projetos precisam ser treinados. Mesmo que, em quatro ocasiões, se procure recrutar gestores profissionais e até mesmo contar com um centro de desenvolvimento de pessoal. Na gestão de projetos é dada ênfase à capacitação para descrever o conteúdo dos bancos de dados com metadados, de grande relevância para melhorar o desempenho das operações nos cenários de *Big Data*.

A presença de especialistas em Bases de dados relacionais garante que a modelagem lógica das tabelas, exibições e índices de dados seja independente das estruturas de armazenamento físico. Essa capacidade se estende às linguagens de consulta usadas nesse cluster para acessar e gerenciar os dados das organizações, tanto estruturados (SQL, seus diferentes padrões e tecnologias ou Neo4J) quanto não estruturados (NoSQL, Dynamo DB), ambos (PostgreSQL) e outros bancos de dados relacionais (IBM DB3, WRDS). Esses programas poderiam ser atendidos no macrocluster Programação e automação de escritórios. Se estão incluídos em Gestão de dados é porque, além da manipulação, os candidatos precisam planejar as entidades e os relacionamentos de BD, pois se requer de alguém que os gerencie ao lado do gestor de dados. Usar as habilidades de gestão de dados no local de trabalho é garantir

o uso do banco de dados, a compreensão dos conceitos de design de *Bases de dados* e o acesso à documentação sobre os processos da organização.

4.1.4 Macrocluster de competências em *Análise de dados e estatísticas*

Se comprova na tabela 12 a importância da experiência em metodologia e ferramentas de análise de dados. Trabalhar como analista de dados requer domínio de habilidades técnicas em matemática, estatística, programação e bancos de dados. Essas ferramentas e técnicas de coleta e análise visam a identificar e extrair informações úteis, modelar os dados e comunicá-los aos executivos e à equipe.

Tabela 12 - Grupos de sinônimos do macrocluster *Análise de dados e estatísticas*.

| Cluster | Ofertas |
|-------------------------------------------|----------------|
| Análise de dados e estatísticas | 51 |
| IA e aprendizado de máquina | 28 |
| Análise e modelagem estatística | 11 |
| Matemática | 6 |
| Big Data e processamento massivo de dados | 4 |

Fonte: dados da pesquisa

Afinidade com números que se associa aos métodos de análise estatística, algoritmos, com base sólida em álgebra e cálculo linear, probabilidade, equações e transformações integradas, Aprendizado de máquina e IA, sempre ligados ao software estatístico. Que se estendem em reduzida presença a seis dos anúncios com um competente desempenho em matemática, destinado a habilitar na gestão da estatística teórica, da lógica formal e da teoria dos conjuntos e que inclui Temas matemáticos e teóricos e Modelagem e otimização matemáticas.

Sem dúvida, os fundamentos teóricos de estatística e matemática junto as habilidades tecnológicas são básicas para a análise de dados (Hardin; Horton, 2017). Então, que tipo de treinamento em matemática seria ideal para um cientista de dados formado em Ciência da Informação com reduzida preparação formal? Com certeza é distinto ao treinamento que seria desejável para um matemático. Mas não é tão estranho quanto aparenta, pois os aspectos quantitativos se têm abordado sempre na CI, como mostrara Brookes (1980), na análise de questões informacionais, em especial sociais, e no fornecimento de uma descrição matemática dos espaços de informação. E cresce o número de aplicações que necessitam das estatísticas de

frequência e das leis logarítmicas (Bawden, 2008) nas pesquisas métricas que aplicam matemática e estatística.

Por outro lado, dentro da natureza multifacetada das habilidades profissionais para a análise de dados, um dos maiores obstáculos tem sido a percepção da necessidade de aprender uma linguagem de programação e de utilizar matemática e estatística avançadas. Uma vez que o desenvolvimento de ferramentas de programação visual disponibilizou software de baixo código e sem código que nivelou o campo da abordagem matemática e permitiu que mais pesquisadores sociais, e é claro da CI, pudessem aplicar análises de ML/AI a conjuntos de dados (Hersch *et al.*, 2023). Voltou se uma destreza relevante o uso de ferramentas de programação intuitivas que facilitem explorar os dados com pouca experiência em programação, criar modelos de dados, avaliar resultados de análises de dados e fazer relatórios de pesquisa de ciência de dados, entre outros (Hersh *et al.*, 2023). De qualquer forma, as ofertas demandam experiência e habilidades no manejo de software para gestão de dados em inteligência de negócios, onde as funções estatísticas do Excel se destacam por suas 24 ocorrências e cuja presença se repete no cluster do Microsoft Office. O mesmo ocorre com os 7 aplicativos de planilhas complexas do Power Bi. O Tableau aparece também quatro vezes. Enquanto outros programas fazem isso uma ou duas vezes, como Stata, Qlick, Iterix, SPSS, Spark, Variety ou aqueles para IA, como Clear forest ou Open Calais.

4.1.5 Macrocluster de competências em Padrões e tecnologias da Web

Padrões e tecnologias da Web são recomendações publicadas pelo World Wide Web Consortium (W3C) que qualquer desenvolvedor web usa para construir sites como HTML/XHTML e Cascading Style Sheets (CSS) como um canal fácil para networking e troca de todos os tipos de dados culturais, comerciais, industriais, científicos e pessoais. Alcançando o campo da ética corporativa e conformidade.

Tabela 13 - Grupos de sinônimos del macrocluster Padrões e tecnologias da Web.

| Cluster | Ofer. |
|-----------------------------------|--------------|
| Padrões e tecnologias da Web | 82 |
| Design e desenvolvimento web | 76 |
| Google e tecnologias relacionadas | 14 |

Fonte: dados da pesquisa

As empresas e organizações estão procurando pessoas com experiência principalmente em tecnologia da web semântica para o gerenciamento de conteúdo que usa padrões semânticos formais baseados em RDF (*Resource Description Framework*), SPARQL (SPARQL Protocol), RDF Query Language e, eventualmente, OWL (*Web Ontology Language*). Bem como na aplicação RDF SKOS (*Simple Knowledge Organization System*) como modelo para representar a estrutura básica e o conteúdo do KOS, e sua implementação no PoolParty Semantic Suite.

A gestão de conteúdo e o desenvolvimento interativo da Web são outras das habilidades mais procuradas, visando à satisfação do usuário (experiência de usuário) e à otimização dos mecanismos de busca (SEO). Dependendo da situação, isso significa ter perícia nas campanhas de mídia social usando plataformas inovadoras de *crowdsourcing*. Também da plataforma de colaboração empresarial MS-SharePoint ou dos editores Figma, WordPress, Tableau e Miro para analisar e visualizar grandes quantidades de dados interativos dedicados ao conhecimento em uma organização ou empresa, resumindo-os e apresentando-os em um formato fácil de entender (Larson; Chang, 2016). A lista é completada por outros sistemas de gestão de conteúdo com menor presença nos anúncios: Draw.io, Drupal, IBM DataStage, Inxight, iServer, Lucidchart, Refinitiv ou Trello. Não falta envolvimento da Microsoft com o Azure, MS-365. MS-Visio, Ms Teams e Ms Fast.

O terceiro cluster é agrupado de acordo com os critérios de atuação nas diferentes plataformas e produtos do Google. Desde o *Google G Suite*, com o típico software de escritório, até o *Google Cloud Platform* (GCP), com todos os aplicativos de desenvolvimento da Web. Há várias ferramentas analíticas cuja gestão é considerada. Por exemplo, o *Looker studio* para converter dados em painéis e relatórios personalizáveis e informativos (Lakshmanan, 2017). *Google Analytics* para rastrear as etapas que os usuários realizam por meio dos dados de diferentes sites e aplicativos. *Google Search Structured Data*, marcação de dados estruturados nas pesquisas do Google. A combinação de dados do *Google Analytics* com dados GNIP das principais redes sociais. E o *Google Sheets*, que poderia muito bem ser classificado como software de gestão de dados.

4.1.6 Macrocluster de competências em Programação e automação de escritórios

A polihierarquia é vista em 15 de seus sinônimos, derivados do fato de ser um software aplicado aos procedimentos de outros macroclusters ou aos dados de campos específicos do conhecimento. O que demonstra a importância das competências tecnológicas, pois seu valor é

aumentado quando somado às habilidades em Padrões e tecnologias da Web ou às relacionadas a software de gestão de dados.

Tabela 14 - Grupos de sinônimos del macrocluster Programação e automação de escritórios.

| Cluster | Ofertas |
|--------------------------------------|---------|
| Microsoft Office | 64 |
| Programação | 35 |
| Suíte Microsoft | 13 |
| Linguagens de programação e consulta | 11 |

Fonte: dados da pesquisa

Os anúncios enfatizam a necessidade de habilidades no uso do conjunto de aplicativos do Microsoft Office para criar, editar, armazenar e transmitir informações por meio de computadores pessoais. Às vezes, as habilidades são exigidas para o pacote completo, mas são especificadas principalmente para o Word e o Excel, este último já mencionado acima para as funções estatísticas na gestão de dados. Além de PowerPoint, Access, VBA e Windows XP.

Já o cluster Programação incorpora conhecimentos especializados em aplicativos e arquitetura de sistemas e interfaces de programação informática, ambientes de nuvem, editores gráficos e estratégias técnicas para o desenvolvimento de produtos. Isso inclui aplicativos de arquitetura, como o TOGAF, a plataforma de programação e computação numérica MTLAB, a multiplataforma de bases de dados relacionais FileMaker Pro, a linguagem de programação orientada a objetos VB.net e até mesmo a aplicação de tecnologias semânticas da Web para o desenvolvimento de soluções empresariais da TopQuadrant.

Isto se aplica ao cluster Linguagens de programação e consulta com habilidades tanto de consulta (R) como de automação e análise de dados estatísticos (Java, Microsoft #, MATLAB, Python e Sparql).

4.1.7 Macrocluster de Habilidades analíticas e criatividade

Por se tratar de um cluster único, os principais sinônimos que o compõem são apresentados a seguir. Algumas de suas habilidades são compatíveis com as do grupo de habilidades *organizacionais, no local de trabalho e interpessoais*, conforme mostrado por suas três tabulações cruzadas polihierárquicas.

Tabela 15 - Grupos de sinônimos do macrocluster Habilidades analíticas e criatividade.

| Sinônimos | F total |
|--------------------------|----------------|
| Atenção aos detalhes | 20 |
| Solução de problemas | 18 |
| Mentalidade colaborativa | 17 |
| Tarefas simultâneas | 15 |
| Criatividade | 15 |
| Habilidades analíticas | 13 |
| Cumprimento de metas | 12 |
| Pensamento crítico | 11 |
| Adaptação a mudanças | 8 |

Fonte: dados da pesquisa

Os termos preferidos de habilidades analíticas e criatividade aparecem em 49 anúncios (58,33%). Por formarem um único cluster, foi dividido em grupos de sinônimos: habilidades analíticas, que permitem observar, investigar e interpretar um assunto para desenvolver ideias e soluções complexas. Como pensamento crítico e habilidades de resolução de problemas para responder aos desenvolvimentos das TIC. O pensamento crítico aumenta a probabilidade de alcançar os resultados desejados quando usado na resolução de problemas encontrados ao longo dos estágios que levam à realização dos objetivos (Halpern, 1998). E ter critérios independentes em encruzilhadas econômicas complexas diante da desinformação e possíveis contradições. Então, ser criativo, capaz de inovar e ter iniciativa são características muito valorizadas em qualquer organização, pois comprovam o comprometimento e a integração de seu pessoal (Khan; Masrek; Nadzar, 2015). Por fim, a vontade de se adaptar à mudança acompanha o desejo de acompanhar a transformação digital. A capacidade tecnológica esteve entre as mais solicitadas nos trabalhos de Choi e Rasmussen (2009), Zhang e Chen (2023) para atender a transformação digital. É a maneira de responder às expectativas dos usuários e seu acesso a uma gama crescente de dados e objetos de conteúdo que geralmente exigem mudanças organizacionais. É abordado a partir de uma atitude de colaboração na consecução de objetivos comuns com colegas de um grupo ou serviço e com atenção aos detalhes - curiosidade intelectual em relação a eventos externos e novos desenvolvimentos.

O desenvolvimento de competências de trabalho consiste, acima de tudo, nas competências de ambiente de trabalho colaborativo que refletem a cooperação entre membros de uma equipa ou entre diferentes departamentos e com parceiros externos. Mais

concretamente, o objetivo é formar equipes multifuncionais cujos membros tenham as capacidades necessárias para resolver situações complexas de forma independente. Nesse cenário, a participação dos profissionais da informação contribui para o alcance dos objetivos do projeto (Semeler, Pinto, Rozados, 2019). Os anúncios mencionam a execução de processos complexos em tarefas simultâneas que exigem atitudes eficientes, éticas e exigentes de dedicação e iniciativa no trabalho. Em vários casos, atribua o trabalho a pessoas com experiência de trabalho em organizações semelhantes e até mesmo caracterize-o como descentralizado e não hierárquico.

4.1.8 *Temas de domínios específicos skills macrocluster*

No cluster Publicação científica e acadêmica sobressaem as habilidades em mudança climática, desenvolvimento sustentável e finanças sustentáveis que enlaçam com as de ciências sociais e comportamentais e ciências socioambientais, que, por sua vez, fazem polihierarquia com Publicação científica e acadêmica.

Tabela 16 - Grupos de sinônimos do macrocluster *Temas de domínios específicos*.

| Cluster | Ofertas |
|------------------------------------|----------------|
| Publicação científica e acadêmica | 26 |
| Saúde e medicina | 10 |
| Atividades artísticas e criativas | 2 |
| Ciências sociais e comportamentais | 2 |
| Temas de domínios específicos | 2 |
| Educação e publicação acadêmica | 1 |

Fonte: dados da pesquisa

É notável o pedido de capacidade de supervisão dos trabalhos escritos por professores e pesquisadores universitários desde levantamentos bibliográficos, aos destinados a periódicos revisados por pares, e até atas de reuniões acadêmicas. Em definitiva os diferentes modelos de publicação na comunicação acadêmica, também os do tipo jornalístico, e o uso de imagens neles, em particular quando se destinam à educação. Além da redação de projetos de pesquisa e desenvolvimento e seus resultados, e até mesmo a aplicação da *Theory of Constraints* (TOC) à solução de problemas científicos.

As capacidades da Publicação científica e acadêmica alcançam à curadoria manual do ChEMBL, banco de dados químicos de moléculas bioativas com propriedades farmacêuticas, do banco de dados *Reactome Biology* e do repositório *Uniprot* de dados de proteínas. Sinônimos

que estão fortemente associados nas ofertas aos do cluster Saúde e medicina que, desde a experiência na gestão de dados de saúde (Beam; Kohane, 2018), se estende à Medicina, diagnósticos médicos, intervenção em crises, Farmácia e produtos farmacêuticos, seguro-saúde, segurança alimentar e até mesmo a questões veterinárias na plataforma de e-learning BHS Wise.

Este macrocluster ainda inclui apelos a perícia em assuntos específicos do campo da religião; experiência na realização de atividades artísticas e criativas usando o editor de gráficos vetoriais Sketch e prática em orientação sexual.

5 CONCLUSÕES

Tendo como fonte os requisitos de competências detectadas nos anúncios de emprego se evidencia um duplo motivo para a CI se envolver na análise e organização de dados:

Que os PI sejam indicados nos anúncios como possíveis candidatos às vagas oferecidas. Para cumprir as tarefas com suficiência, os anunciantes buscam integrar nas suas equipes a pessoal com uma ampla variedade de formação, mas cuja qualificação seja adequada para responder à gestão dos mais diversos dados e informações. O PI, enquanto analista de dados e conteúdo, se associa a outros especialistas, notadamente de ciência de dados, STEM, administração e gestão, e comunicação, para compartilhar seus conhecimentos e capacidades. Ao tempo que para interpretar procedimentos e produtos em todas as divisões da organização. Há uma forte tendência de trabalhar com colegas alheios ao âmbito formativo de proveniência, que significa atuar num campo interdisciplinar desafiador.

Devido à destreza nas técnicas de recuperação da informação e de organização do conhecimento exigidas pelos anúncios. O PI contribui com seu domínio técnico na etiquetagem com metadados para identificar a origem dos dados e a frequência de atualização, mantê-los em suas séries e preservá-los na fase de análise. É a maneira pela qual se colocam os dados no contexto certo que dá significado e coerência à interpretação de seu conteúdo. Já na gestão assiste desde a capacidade de representação coerente dos dados relativos aos serviços e produtos das organizações e empresas em diferentes ambientes de atuação. Em especial nos *Big Data*, cujo problema não é apenas a magnitude, mas a descontextualização. Em resposta, uma abordagem semântica das taxonomias pode integrar os resultados da análise desses dados aos sistemas de informação e à tomada de decisões segundo o conhecimento aportado. Isso também garante a habilitação para a interoperabilidade de vocabulários e esquemas de metadados a fim de formalizar seus modelos e lograr a conformidade das taxonomias com a análise de dados.

Por causa da natureza empresarial e comercial dos empregos oferecidos, as competências em Gestão de negócios e empresas é um dos macroclusters com maior número de sinônimos. Nele se combinam habilidades para a análise das ferramentas e infraestruturas de dados com as melhores práticas para o desempenho dos processos operativos, dos recursos humanos e materiais e das estratégias técnicas de negócios, algumas delas bem específicas. Com extensão às aptidões de negócio desde fundamentos econômicos e de conhecimento do mercado que implicam flexibilidade para novas iniciativas como a experiência do usuário para comércio eletrônico. As habilidades interpessoais também são de interesse quando se trata de conectar as unidades envolvidas em uma operação e os clientes. As capacitações em comunicação, sobretudo para transferir o conhecimento e a imagem corporativa, fazem parte desse grupo e são inseparáveis das habilidades de expressão linguística e do domínio das técnicas multimídia.

A maioria das qualidades exigidas nos anúncios corrobora os resultados de análises anteriores para os profissionais de bibliotecas e informação, portanto, não representam uma grande mudança em suas propostas, embora representem uma consolidação. É o caso das habilidades gerenciais, operacionais e de liderança que propiciam a realização de projetos bem-sucedidos e que sempre caracterizaram seu desempenho profissional, além daquelas relativas ao relacionamento com os usuários. Ao mesmo tempo, as competências em informação favorecem as mudanças institucionais exigidas pelo contexto organizativo.

Com relação ao acúmulo contínuo dos dados e seu desempenho na administração, custódia e distribuição de objetos digitais os anúncios requerem aptidões técnicas destinadas a tornar os dados rastreáveis, acessíveis, interoperáveis e reutilizáveis por um longo período. A missão da Gestão de dados é responder às expectativas e aos requisitos de empresas e instituições, e até mesmo de agências de financiamento, no caso de pesquisadores, consoante com a tradição dos documentalistas e bibliotecários acadêmicos de terem habilidades gerenciais e organizativas. Nesse grupo, a contribuição da CI para o tratamento de dados é determinada pela hibridização com que os dados, as informações ou o conhecimento são explorados, pois, além de serem colocados em prática em um campo específico, envolvem competências gerenciais que, por sua vez, resultam impossíveis sem as habilidades relacionadas a utilização dos bancos de dados e das TIC. No entanto, poucas instituições educativas os preparam para isso, assim sendo, os solicitantes precisam que seus estudos universitários incluam a habilitação nesses objetivos ou ser treinados no local de trabalho.

As empresas e organizações multinacionais reconhecem as competências dos especialistas em informação na análise e na gestão de dados, conforme demonstrado pelo fato de que, entre os macroclusters de competências, as competências em Gestão da informação, biblioteconomia e organização do conhecimento aparecem em terceiro lugar pela frequência simples de sinônimos. Essa confiança se baseia novamente no conhecimento sobre padrões, formatos e métodos normalizados de descrição, apresentação e transmissão. Portanto, os anúncios requerem experiência e capacidade de pesquisar e integrar dados de diferentes fontes a fim de melhorar a taxonomia de produtos e serviços, para atender às necessidades atuais e futuras dos clientes e da empresa. Com a oportunidades de criar vocabulários semânticos adequados aos big data, tornando-os legíveis para os usuários e para os cientistas de dados. Não há muitos trabalhos acadêmicos disponíveis que forneçam uma taxonomia para as técnicas de análise de Big Data. Portanto, aqui os anúncios estão à frente da pesquisa sobre arquitetura de dados semânticos.

Enquanto o arquiteto e o cientista de dados são perfis avançados no trabalho com dados e nas capacitações em computação, o perfil do analista é mais compatível com uma formação universitária mais flexível. Vimos nos resultados a importância das habilidades de comunicação e linguagem na fase de interpretação dos dados e de dar a eles coerência discursiva em relação aos departamentos e equipes da organização e aos usuários. A solicitação dos anunciantes por pessoas com formação universitária em Ciências Sociais e Humanas se baseia nesse argumento. É perfeitamente possível que um graduado em CI possa trabalhar nessas equipes, mas isso não é gratuito. Pois certamente precisará de boas habilidades estatísticas e de TIC do tipo já fornecido por muitos currículos com múltiplas disciplinas atuais e práticas, se não for adquirido por meio de cursos de mestrado ou especialização, ou mesmo por meio de treinamento no trabalho e com a experiência profissional.

As plataformas de análise fizeram uma contribuição significativa para o processamento dos dados por meio de suas interfaces intuitivas. As plataformas deram eficácia representativa às necessárias planilhas mediante visualizações claras, interativas e fáceis de compartilhar. De uma forma que otimiza o poder analítico dos dados contidos nas planilhas.

Os anúncios abrangem o campo das TIC de forma ampla, pois afeta vários dos macroclusters nos quais os termos preferidos estão organizados. São aqueles que envolvem habilidades técnicas e alto conhecimento de software para aplicação em procedimentos e processos em constante mudança. As habilidades digitais estão muito próximas das habilidades de navegação na web, sua estrutura e ferramentas de busca e redes sociais. A obtenção dessas

competências é, conforme observado acima, abordada pelas diversas disciplinas de tecnologia ensinadas em muitos cursos de CI. Portanto, o treinamento tecnológico é inerente aos profissionais da informação. Não apenas por ser um componente inseparável da CI, mas também por ter se fortalecido tanto desde a transformação digital que teve um impacto radical nas competências, com o consequente reflexo nas mudanças curriculares em nível de graduação e pós-graduação e até mesmo na própria denominação dos cursos e perfis profissionais.

Isto posto, enfatiza-se que a análise dos anúncios de emprego é uma ferramenta que pode ser usada de diversas maneiras para caracterizar as profissões da informação e sustentar decisões sobre seu planejamento curricular e, é claro, sobre as necessidades de treinamento no trabalho e outras questões relativas a recursos humanos.

Entretanto, ainda há muito a ser explorado em termos das competências que os profissionais da informação precisam ter para serem proficientes nas tarefas de análise e gestão de dados. Junto com as implicações que isso terá para a identidade profissional da profissão e de seus conselhos e associações.

REFERÊNCIAS

- AGARWAL, R.; DHAR, V. Editorial-Big Data, Data Science, and Analytics: The Opportunity and Challenge for IS Research. **Information Systems Research**, v. 25, n. 3, p. 443–448, 2014. Disponível em: <https://doi.org/10.1287/isre.2014.0546>. Acesso em: 01 jun. 2025.
- ANSI/NISO. Z39.19-2005. **Guidelines for the Construction, Format and Management of Monolingual Controlled Vocabularies**. Bethesda, Maryland: NISO Press, 2005.
- BAWDEN, David. Smoother pebbles and the shoulders of giants: the developing foundations of information science. **Journal of Information Science**, v. 34, n. 4, p. 415-426, 2008. Disponível em: <https://doi.org/10.1177/0165551508089717>. Acesso em: 01 jun. 2025.
- BEAM Andrew L, Kohane Isaac S. Big Data and Machine Learning in Health Care. **Journal of the American Medical Association**, v. 319. n. 13, p. 1317–1318, 2018. Disponível em: <https://doi.org/10.1001/jama.2017.18391>. Acesso em: 08 maio 2025.
- BEREJO, Antonio. Theoretical foundations of Library and Information Science. An epistemological and methodological approach. Balisages. **La revue de recherche de l'Enssib**, v. 1, 2019. Disponível em: <https://publications-prairial.fr/balisages/index.php?id=85>. Acesso em: 09 abr. 2025.
- BERELSON, Bernard. **Content Analysis in Communication Research**. New York: Free Press, 1952.
- BOLAÑOS-MEJÍAS, Carmen; BUENO DE LA FUENTE, Gema; ALVES, Fernanda Melo. O protagonismo da informação-documentação na cooperação ao desenvolvimento: os recursos educativos abertos nos processos de e-learning. **Informação & Sociedade: Estudos**, v. 23, n. 1, p. 105-115, 2013. Disponível em: <https://periodicos.ufpb.br/index.php/ies/article/view/15120/9269>. Acesso em: 01 jun. 2025.
- BONATTI, Piero; DECKER, Stefan; POLLERES, Axel; PRESUTTI, Valentina. Knowledge graphs: New directions for knowledge representation on the semantic web (dagstuhl seminar 18371). **Dagstuhl reports**, v. 8, n. 9, p. 29-11, 2019. Disponível em: <https://doi.org/10.4230/DagRep.8.9.29>. Acesso em: 06 maio 2025.
- BROOKES, Bertram. C. The foundations of information science, Part II. Quantitative aspects: classes of things and the challenge of human individuality. **Journal of Information Science**, v. 2, n. 5, p. 209-21, 1980. <https://doi.org/10.1177/016555158000200502>. Acesso em: 08 maio 2025.
- CHOI, Youngok; RASMUSSEN, Edie. What qualifications and skills are important for digital librarian positions in academic libraries? A job advertisement analysis. **Journal of Academic Librarianship**, v. 35, p.457-467, 2009. Disponível em: <https://doi.org/10.1016/j.acalib.2009.06.003>. Acesso em: 08 maio 2025.
- CRISCI, Jorge; LÓPEZ-ARMENGOL, María Fernanda. (1983). **Introducción a la teoría y práctica de la taxonomía numérica**. Washington: OEA. Disponível em: <https://www.researchgate.net/publication/50326234>. Acesso em: 24 maio 2025.

EROZKAN, Atilgan. The effect of communication skills and interpersonal problem-solving skills on social self-efficacy. **Educational Sciences: Theory and Practice**, v. 13, n. 2, p. 739-745, 2013. Disponível em: <https://files.eric.ed.gov/fulltext/EJ1017303.pdf>. Acesso em: 18 abr. 2025.

FEDERER, Isa. Defining Data Librarianship: A Survey of Competencies, Skills, and Training. **Journal of Medical Librarianship**, v. 106, n. 3, p. 294-303, 2018. Disponível em: <https://doi.org/10.5195/jmla.2018.306>. Acesso em: 09 jun. 2025.

GANI, Abdullah; SIDDIQA, Aisha; SHAMSHIRBAND, Shahaboddin; HANUM, Fariza. A survey on indexing techniques for big data: Taxonomy and performance evaluation. **Knowledge and information systems**, v. 46, n. 2, p. 241-284, 2015. Disponível em: <https://doi.org/10.1007/s10115-015-0830-y>. Acesso em: 19 jun. 2025.

GIVEN, Lisa M.; OLSON, Hope A. Knowledge organization in research: A conceptual model for organizing data. **Library & Information Science Research**, v. 25, n. 2, p. 157-176, 2003. Disponível em: [https://doi.org/10.1016/S0740-8188\(03\)00005-7](https://doi.org/10.1016/S0740-8188(03)00005-7). Acesso em: 09 jun. 2025.

GOULDING, Anne; BROMHAM, Beth; HANNABUSS, Stuart; CRAMER, Duncan. Supply and demand the workforce needs of library and information services and personal qualities of new professionals. **Journal of Librarianship and Information Science**, v. 31, p. 212-223, 1999. Disponível em: <https://doi.org/10.1177/096100069903100405>. Acesso em: 19 jun. 2025.

GRIFFEY, Jason (ed.). **Artificial Intelligence and Machine Learning in Libraries**. Chicago: ALA, Library Technology Reports, 2019.

HAIR, Josep; BLACK, William; BABIN, Barry; ANDERSON, Rolf; TATHAM, Ronald. **Análise multivariada de dados**. 6.ed. Bookman, 2009.

HALPERN, Diane. Teaching critical thinking for transfer across domains: disposition, skills, structure training, and metacognitive monitoring. **American Psychologist**, v. 53, n. 4, p. 449-455, 1998. Disponível em: <https://doi.org/10.1037/0003-066X.53.4.449>. Acesso em: 02 jun. 2025.

HARDIN, Johanna S.; HORTON, Nicholas J. Ensuring that mathematics is relevant in a world of data science. **Notices of the AMs**, v. 64, n. 9, p. 986-990, 2017. Disponível em: <https://www.ams.org/publications/journals/notices/201709/rmoti-p986.pdf>. Acesso em: 19 jun. 2025.

HENDERSON, Margaret. **Gestão de dados: a pratical guide for librarians**. Lanham: Rowman & Littlefield Publishers, 2017.

HERNDON, Joel. **Data Science in the Library: Tools and Strategies for Supporting Data-Driven Research and Instruction**. London: Facet Publishing, 2021.

HERSH, William R.; HOYT, R. E.; CHAMBERLIN, S.; ANCKER, J. S.; GUPTA, A.; BORLAWSKY-PAYNE, T. B. Beyond mathematics, statistics, and programming: data science, machine learning, and artificial intelligence competencies and curricula for clinicians, informaticians, science journalists, and researchers. **Health Systems**, v. 12, n. 3, p. 255-263, 2023. Disponível em: <https://doi.org/10.1080/20476965.2023.2237745>. Acesso em: 18 jun. 2025.

HUSAMALDIN, Laden; Saeed, Nagham. Big data analytics correlation taxonomy. **Information**, v. 11, n. 1, p. 17, 2019. Disponível em: <https://doi.org/10.3390/info11010017>. Acesso em: 19 jun. 2025.

KELLAM, Linda; THOMPSON, Kristi. **Databrarianship: The academic data librarian in theory and practice**. Chicago: American Library Association, 2016.

KENNAM, Mary Anne. Gestão de dados: Knowledge and skills required in research, scientific and technical organisations. In: IFLA World Library and Information Congress: 82nd IFLA General Conference and Assembly. International Federation of Library Associations and Institutions. **Proceedings...** p. 1-10, 2016. Disponível em: https://researchoutput.csu.edu.au/ws/portalfiles/portal/20227185/9761721_Published_paper.pdf. Acesso em: 20 jun. 2025.

KHAN, Asad; MASREK, Noorman; Nadzar, Fuziah. Analysis of competencies, job satisfaction and organizational commitment as indicators of job performance: A conceptual framework. **Education for Information**, v. 31, n. 3, p. 125-141, 2015. Disponível em: <https://doi.org/10.3233/efi-150954>. Acesso em: 03 jun. 2025.

KHAN, Hammad Rauf; DU, Yunfei. What is a Data Librarian? A Content Analysis of Job Advertisements for Data Librarians in the United States Academic Libraries. IFLA WLIC 2018 – Kuala Lumpur, Malaysia – Transform Libraries, Transform Societies. Session 139 - Education and Training, Science and Technology, and Continuing Professional Development and Workplace Learning. **Proceedings...** 2018. Disponível em: <https://library.ifla.org/id/eprint/2255/>. Acesso em: 03 jun. 2025.

KINKUS, Jane. Gestão de projetos skills: A literature review and content analysis of librarian position announcements. **College & Research Libraries**, v. 68, n. 4, p. 352-363, 2007. Disponível em: <https://doi.org/10.5860/crl.68.4.352>. Acesso em: 03 jun. 2025.

KONOPÁSEK, Zdeněk. Making thinking visible with Atlas.ti: Computer assisted qualitative analysis as textual practices. **Historical Social Research Supplement**, v. 19, p. 276-298, 2007. GESIS - Leibniz Institute for the Social Sciences. Disponível em: <https://www.jstor.org/sTabela/40981082>. Acesso em: 08 jun. 2025.

KRIPPENDORF, Klaus. **Content analysis: An introduction to its methodology**. Sage publications, 2018.

LAKSHMANAN, Valliappa. **Data Science on the Google Cloud Platform**. O'Reilly Media, 2017.

LARSON, Dianne; CHANG, Victor. A review and future Direction of agile, business inteligente, analytics, and data Science. **Internationale Journal of Information Management**, v. 36, n. 5, p. 700-710, 2016. Disponível em: <https://doi.org/10.1016/j.ijinfomgt.2016.04.013>. Acesso em: 20 maio 2025.

LOMA, Jane. How to use Ontologies to Orlock the full potential of hour Scientific Data – Part 1. Sibite. Blog. 2019. Disponível em: <https://www.scibite.com/news/how-to-use-ontologies-to-unlock-the-full-potential-of-your-scientific-data-part-1/>. Acesso em: 12 jun. 2025.

MAYR, Philipp; TUDHOPE, Douglas; CLARKE DEXTRE, Stella; ZENG LEI, Marcia; LIN, Xia. Recent applications of Knowledge Organization Systems: introduction to a special issue. **International Journal on Digital Libraries**, v. 17, p. 1-4, 2016. Disponível em: <https://doi.org/10.1007/s00799-015-0167-x>. Acesso em: 20 maio 2025.

MEGANCK, Shana; Smith, Joshua; Guidry, Jeanine. The skills required for entry-level public relations: An analysis of skills required in 1,000 PR job ads. **Public Relations Review**, v. 46, n. 5, p. 101973, 2020. Disponível em: <https://doi.org/10.1016/j.pubrev.2020.101973>. Acesso em: 10 maio 2025.

MOREIRO-GONZÁLEZ, José Antonio; Paletta, Francisco C.; Pastor-Sánchez, Juan Antonio; Coelho-Neves, Bárbara. The Skills of Information Professionals for the Analysis and Organization of Data. Conceptual Proposal and Application of OpenAI in Job Advertisements on the Web. **Profesional de la información**, v. 33, n. 5, e330501, 2024. Disponível em: <https://doi.org/10.3145/epi.2024.0501>. Acesso em: 12 jun. 2025.

MOREIRO-GONZÁLEZ, José Antonio, VERGUEIRO, Waldomiro. Ofertas de trabalho na Web para os profissionais brasileiros da informação-documentação: análise das competências e habilidades exigidas pelas empresas e instituições. **Perspectivas em ciência da informação**, v. 17, p. 231-250, 2012. Disponível em: <https://doi.org/10.1590/S1413-99362012000100013>. Acesso em: 20 jun. 2025.

NICKERSON, Robert; VARSHNEY, Upkar; MUNTERMANN, Jann. Method for Taxonomy Development and Its Application in Information Systems. **European Journal of Information Systems**, v. 22, n. 3, p. 336-359, 2013. Disponível em: <https://doi.org/10.1057/ejis.2012.26>. Acesso em: 20 jun. 2025.

NOY, N.; GAO, Y; JAIN, A.; NARAYANAN, A; Patterson; TAYLOR, J. Industry-scale knowledge graphs: lessons and challenges. **ACM Queue**, v. 17, n. 2, p. 48-75, 2019. Disponível em: <https://doi.org/10.1145/3329781.3332266>. Acesso em: 18 maio 2025.

OLIVER, Jeffrey; KOLLEN, Christine; HICKSON, Benjamin; RIOS, Fernando. (2019). Data science support at the academic library. *Journal of Library Administration*, v. 59, n. 3, pp. 241-257. Disponível em: <https://doi.org/10.1080/01930826.2019.1583015>. Acesso em: 12 jun. 2025.

PALETTA, Francisco-Carlos; MOREIRO-GONZÁLEZ, José-Antonio. A transformação digital e os impactos no mercado de trabalho estudo dos anúncios de emprego na web para profissionais da Informação no setor privado. **Information research**, v. 26, n. 3, 2021. Disponível em: <https://doi.org/10.47989/irpaper904>. Acesso em: 20 jun. 2025.

PALETTA, Francisco-Carlos; PASTOR-SÁNCHEZ, Juan-Antonio; MOREIRO-GONZÁLEZ, José-Antonio. Competências e habilidades digitais requeridas aos profissionais da informação nos anúncios brasileiros de emprego na web. **Profesional de la información**, v. 30, n. 1, e300108, 2021. Disponível em: <https://doi.org/10.3145/epi.2021.ene.08>. Acesso em: 18 jun. 2025.

PAPADIMITRIOU, Panagiotis; TSAPARAS, Panayiotis; FUXMAN, Ariel; GETOOR, Lise. TACI: Taxonomy-Aware Catalog Integration. **IEEE Transactions on Knowledge and Data Engineering**, v. 25, n. 7, p. 1643-1655, 2013. Disponível em: <https://doi.org/10.1109/TKDE.2012.54>. Acesso em: 05 jun. 2025.

SEADLE, Michael; HAVELKA, Stefanie. Information science: Why it is not data science. **Data and Information Management**, v. 7, 100027, 2023. Disponível em: <https://doi.org/10.1016/j.dim.2023.100027>. Acesso em: 20 jun. 2025.

SEMELER, Alexandre; PINTO, Adilson; KOTAY, Tibor; DIAS, Thiago; OLIVEIRA, Arthur; MOREIRO-GONZÁLEZ, José A.; ROZADOS, Helen-Beatriz. Algorithmic Literacy: Generative Artificial Intelligence Technologies for Data Librarians. **EAI Endorsed Transactions on Scalable Information Systems**, v. 11, n. 2, 2024. Disponível em: <http://dx.doi.org/10.4108/eetsis.4067>. Acesso em: 04 jun. 2025.

SEMELER, Alexandre; PINTO, Adilson Luis; ROZADOS, Helen-Beatriz. Data science in data librarianship: Core competencies of a data librarian. **Journal of librarianship and information science**, v. 51, n. 3, p. 771-780, 2019. Disponível em: <https://doi.org/10.1177/0961000617742465>. Acesso em: 18 jun. 2025.

STUART, David. **Practical Data Science for Information Professionals**. London: Facet Publishing, 2020.

ZHANG, Jin; Chen, Jianyao. Skill analysis of library and information science professionals. **Journal of Librarianship and Information Science**, v. 57, n. 1, 2023. Disponível em: <https://doi.org/10.1177/09610006231207656>. Acesso em: 07 maio 2025.