

ORGANIZAÇÃO DO CONHECIMENTO E INTELIGÊNCIA ARTIFICIAL: possíveis contribuições para a Arquivologia

KNOWLEDGE ORGANIZATION AND ARTIFICIAL INTELLIGENCE: possible contributions
to archival science

Jayron Viana dos Santos | Sonia Maria Troitiño Rodriguez

<https://doi.org/10.21747/21836671/pagesp2025a14>

Resumo: O presente estudo tem como objetivo analisar como as ferramentas de Inteligência Artificial (IA) podem contribuir para o trabalho arquivístico sob a perspectiva dos Sistemas de Organização do Conhecimento (SOC). Busca-se compreender de que forma essas tecnologias organizam e utilizam conhecimentos para gerar subprodutos informacionais e apoiar os processos de organização, tratamento e recuperação da informação em arquivos. A pesquisa, de natureza qualitativa, descritiva e exploratória, baseia-se em revisão bibliográfica e na análise de estudos teóricos e aplicações práticas relacionadas à IA no campo da Arquivologia. O foco recai especialmente nos métodos de aprendizagem, estruturação e organização de dados e informações por meio de *machine learning* (aprendizado de máquina) e *deep learning* (aprendizado profundo). Espera-se contribuir para ampliar as discussões e debates, destacando as possíveis contribuições da IA na melhoria da organização e sistematização do conhecimento em arquivos, possibilitando novas práticas em processos arquivísticos.

Palavras-chave: Arquivologia; Inteligência artificial; Sistemas de organização do conhecimento.

Abstract: This study aims to analyse how Artificial Intelligence (AI) tools can contribute to archival work from the perspective of Knowledge Organisation Systems (KOS). The aim is to understand how these technologies organise and use knowledge to generate information by-products and support the processes of organising, processing and retrieving information in archives. The research, which is qualitative, descriptive and exploratory in nature, is based on a literature review and analysis of theoretical studies and practical applications related to AI in the field of archives. The focus is especially on methods for learning, structuring and organising data and information using machine learning and deep learning. It is hoped to contribute to broadening discussions and debates, highlighting the possible contributions of AI to improving the organisation and systematisation of knowledge in archives, enabling new practices in archival processes.

Keywords: Archival Science; Artificial Intelligence; Knowledge organization systems.

1. Introdução

O crescimento exponencial do volume de informações e o surgimento de novas formas de registrá-las, especialmente a partir da segunda metade do século XX, intensificaram a complexidade do processo de organização do conhecimento, especialmente no que diz respeito à representação e ao acesso à informação.

Recentemente, algumas pesquisas no campo da Ciência da Informação vêm se dedicando a acompanhar os avanços da Inteligência Artificial (IA) e suas possíveis contribuições em diversas áreas e circunstâncias de uso, ofertando novos pontos de vista e propondo melhorias. Entre os possíveis beneficiários dos avanços da IA, está a Arquivologia, área de estudo que se dedica aos arquivos e seus documentos, nas dimensões administrativa, social

e histórica, dessa forma, contribuindo com a CI, especialmente no que tange aos sistemas de classificação e recuperação de documentos e informações.

A partir da organização desse conhecimento, criam-se recursos que apresentam interpretações estruturadas em esquemas de representação, chamados de Sistemas de Organização do Conhecimento (SOC), que para Dahlberg (2006) destinam-se a facilitar a recuperação de informações e suas conexões.

Sob essa ótica e em relação às IA, o desenvolvimento crescente de tecnologias e ferramentas têm proporcionado agilidade e celeridade a inúmeras atividades por meio de processos automatizados inteligentes, incidindo de maneira significativa no cotidiano do fazer arquivístico. Dessa maneira, contribuem indubitavelmente para a organização, tratamento, gestão e recuperação de documentos e de seus contextos de produção e uso ao longo do tempo.

Diante desse cenário, a questão norteadora desta pesquisa foi: quais as possíveis contribuições das ferramentas e soluções de IA ao processo de organização do conhecimento em arquivos no escopo da Ciência da Informação? Nesse intuito, o objetivo geral foi analisar como as ferramentas de IA podem contribuir para o trabalho arquivístico a partir da perspectiva dos SOC - partindo da premissa de que essas soluções ou recursos de IA adquirem conhecimentos para gerar subprodutos informacionais, sistematizando, tratando e recuperando a informação em arquivos.

Especificamente, objetivou-se buscar: (i) refletir sobre o papel das ferramentas de IA no processo de construção e organização do conhecimento em SOC; (ii) perceber a interação entre as ferramentas de IA e Organização do Conhecimento, com destaque para o emprego em atividades e processos arquivísticos.

Para tanto, foi realizada uma pesquisa bibliográfica, de natureza qualitativa do tipo descritiva e exploratória, utilizando bases de dados referenciais e especializadas, além da busca e descrição de ferramentas de IA com possibilidades de uso em alguns processos arquivísticos, prescindindo para tanto a organização e sistematização do conhecimento em arquivos, considerando os métodos de aprendizagem, estrutura e organização de dados e informações de *machine learning* (aprendizado de máquina) e *deep learning* (aprendizado profundo).

2. Procedimentos metodológicos

As bases de dados consultadas para o desenvolvimento da pesquisa bibliográfica foram a BRAPCI (Base de Dados Referenciais de Artigos de Periódicos em Ciência da Informação), a BDA (Bases de Dados em Arquivística) e o Repositório Institucional UNESP (Universidade Estadual Paulista, Brasil), utilizando termos de busca e palavras-chave com operadores booleanos: “inteligência artificial AND arquivologia”; “Organização do conhecimento AND arquivologia”, “Organização do conhecimento AND inteligência artificial” e “ferramentas de IA”. Em complemento, para análise das fontes recuperadas, procedeu-se à leitura do resumo e palavras-chave para escolha das fontes pertinentes, priorizando textos com relação entre as temáticas abordadas.

Posteriormente, realizou-se uma busca específica no material analisado acerca de ferramentas já utilizadas em alguns processos de organização do conhecimento em sistemas de informação, mais precisamente, em ambientes de arquivo, propiciando uma maior relação e fundamentação do estudo.

Assim, esta pesquisa é caracterizada ainda como qualitativa, que na perspectiva de Creswell é fundamentalmente interpretativa, ou seja, o pesquisador interpreta os dados partindo de uma visão ampla e holística dos fenômenos. “[...] Isso explica por que estudos de pesquisa qualitativa aparecem como visões amplas em vez de microanálises. Quanto mais complexa, interativa e abrangente a narrativa, melhor o estudo qualitativo” (CRESWELL, 2007:187).

Aliado a isso, também recorreu-se à pesquisa exploratória e descritiva, onde a pesquisa exploratória possibilita que o pesquisador analise os dados qualitativos de maneira sistemática, promovendo uma compreensão ou interpretação detalhada do fenômeno em estudo (LÖSH, RAMBO e FERREIRA, 2023). E a pesquisa descritiva, que observa, registra, analisa e organiza os dados sem manipulá-los, ou seja, sem que haja interferência direta do pesquisador (PRODANOV e FREITAS, 2013).

3. Organização do Conhecimento e Inteligência Artificial

A Organização do Conhecimento (OC) é um campo interdisciplinar que se preocupa com os processos de estruturação, classificação e representação do conhecimento, tendo em vista facilitar seu acesso e recuperação. Nesse contexto, destacam-se duas perspectivas principais: a visão adotada pela Ciência da Informação e a noção de conhecimento objetivo ou objetivado, que é registrado em documentos.

A Ciência da Informação, conforme apontado por Hjørland (2008), enxerga o campo da organização do conhecimento como uma atividade fundamental para viabilizar a mediação entre o usuário e a informação. Nesse sentido, desenvolvem-se sistemas e metodologias que auxiliem na representação e na recuperação de informações relevantes, valorizando o papel das linguagens documentárias, tesouros e ontologias como instrumentos para a organização sistemática de informações em ambientes digitais e tradicionais.

A noção de conhecimento objetivo, conforme discutida por autores como Duranti (1993) e Buckland (1991), concentra-se no registro material do conhecimento, ou seja, nos documentos e na informação neles contida. Desse modo, os documentos são mais do que meros suportes; eles são artefatos culturais que encapsulam o conhecimento humano de forma tangível, possibilitando sua transmissão, conservação e estudo. Essa visão enfatiza que o tratamento documental é uma atividade central para garantir a preservação e o acesso ao conhecimento acumulado pela humanidade.

De maneira geral, ambas as perspectivas são complementares e fundamentais para a organização do conhecimento. Enquanto a Ciência da Informação se preocupa com os processos e ferramentas que facilitam o acesso e o uso do conhecimento, a noção de conhecimento objetivo foca na importância do registro documental como base para o trabalho organizacional. Essa complementaridade reflete-se na prática, onde a organização do conhecimento busca equilibrar essas dimensões para atender às demandas de um mundo cada vez mais dependente de informação estruturada e acessível.

A International Society for Knowledge Organization (ISKO) foi criada como associação profissional dedicada ao estudo e promoção da organização do conhecimento. Fundada em 1989 por Ingetraut Dahlberg, a ISKO reúne acadêmicos e profissionais de diversas disciplinas, incluindo Ciência da informação, Filosofia, Linguística e Ciência da Computação, com o objetivo de avançar na teoria e prática da organização do conhecimento.

O investimento científico dos pesquisadores na construção de bases e no campo da Ciência da Informação contribuiu para o surgimento de uma área de pesquisa e, conseqüentemente, de uma disciplina dedicada à promoção de estudos, desenvolvimento e aplicações de sistemas conceituais para a organização do conhecimento.

Essa área também abrange a análise dos aspectos filosóficos e semânticos relacionados à organização conceitual e à dinâmica do conhecimento, onde os aspectos filosóficos incluem questões sobre a natureza do conhecimento e como ele é representado de forma objetiva e compartilhável, enquanto a semântica explora as relações entre conceitos e significados para criar estruturas de classificação que reflitam diferentes contextos e interpretações.

Nota-se uma evolução significativa no campo da OC, que passou de seu objetivo inicial de desenvolver um sistema universal de organização do conhecimento para as abordagens contemporâneas, reconhecendo a pluralidade de perspectivas e incorporando múltiplos propósitos, promovendo ações colaborativas e representativas, especialmente em ambientes digitais que demandam sistemas flexíveis e inclusivos (SANTOS, NEVES e SOUZA, 2019).

De maneira geral, a OC tem como propósito sistematizar e organizar o conhecimento, visando facilitar a recuperação da informação. Além disso, é reconhecida como um campo que desempenha funções importantes de comunicação, mediação e transferência de saberes, contribuindo para os processos de ensino e aprendizagem e apoiando as atividades de pesquisa científica (MORAES, 2017; HJØRLAND, 2008).

Na perspectiva da OC, organizar e sistematizar o conhecimento são atividades que se complementam, mas possuem diferenças fundamentais. Organizar o conhecimento refere-se ao processo de classificar, categorizar e estruturar informações para que elas possam ser acessadas e recuperadas de maneira eficiente. Esse processo envolve a utilização de ferramentas como taxonomias, ontologias e sistemas de classificação, que facilitam a navegação e o uso do conhecimento armazenado. Por outro lado, sistematizar o conhecimento vai além da organização, pois abrange a criação de uma estrutura lógica e coerente que conecta diferentes elementos do conhecimento, proporcionando uma visão integrada e compreensível de um determinado campo ou tópico.

Dahlberg (1995) e Hjørland (2008) contribuem para essa diferenciação. Dahlberg enfatiza que organizar o conhecimento é essencial para criar estruturas classificatórias que viabilizem a localização e o acesso à informação, enquanto Hjørland destaca que a sistematização do conhecimento requer a construção de relações semânticas que reflitam a compreensão contextual e o significado profundo dos conceitos. Assim, organizar é visto como um processo prático e funcional, enquanto sistematizar envolve uma dimensão mais teórica e analítica, integrando conhecimento de maneira mais holística.

Para Dahlberg (2006), a OC é um campo de estudos dedicado à estruturação e sistematização dos conceitos com base em suas características, definidas como elementos

herdados do objeto. Esse domínio de estudos abrange a aplicação de conceitos e suas classes, organizados conforme a atribuição de valores e os conteúdos referentes aos objetos ou assuntos.

Dessa OC, surgem ferramentas que oferecem uma interpretação estruturada e organizada dos objetos, conhecidas como sistemas de organização do conhecimento.

Dahlberg (2013, *apud* OHLY, 2023) apresentou dez desideratas que contribuíram significativamente para o amadurecimento da OC como campo de estudo. A primeira delas enfatiza a importância de reconhecer as unidades de um sistema de ordenação como conceitos de conhecimento, utilizando suas características para estabelecer uma ordem estruturada.

O conceito de "sistema de ordenação de unidades/conceito de conhecimento", conforme proposto por Dahlberg (2006), constitui a base teórica para o desenvolvimento dos SOC. Esses sistemas são projetados como recursos que apresentam interpretações estruturadas por meio de esquemas de representação, com o objetivo primordial de organizar a informação e o conhecimento, facilitando sua recuperação.

Essa perspectiva ressalta a importância de uma estruturação lógica e semântica que permita a interoperabilidade entre diferentes domínios de conhecimento. De acordo com Hjørland (2016), os SOC desempenham um papel essencial na mediação do acesso à informação, uma vez que possibilitam a criação de vínculos significativos entre os conceitos, promovendo um entendimento mais aprofundado e contextualizado. Nesse sentido, os SOC transcendem a simples categorização, tornando-se ferramentas indispensáveis para a gestão do conhecimento em ambientes complexos e dinâmicos.

A análise de Gnoli (2020) reforça essa ideia ao destacar que os SOC, ao utilizarem esquemas de classificação, taxonomias, ontologias e tesouros, permitem não apenas a organização, mas também a evolução dos sistemas de conhecimento. Essa evolução ocorre à medida que novos conceitos são integrados, ajustados ou reestruturados, em resposta às mudanças nos campos científicos e tecnológicos.

Além disso, a eficácia dos SOC está intrinsecamente ligada à clareza e coerência das representações conceituais que eles empregam. Assim, os SOC não apenas facilitam a recuperação da informação, mas também contribuem para a construção de novos conhecimentos, ao fornecer uma base sólida para o raciocínio e a análise.

Portanto, os SOC, ancorados no sistema de ordenação de unidades/conceito de conhecimento delineado por Dahlberg, representam um avanço significativo na organização e gestão da informação. Eles atendem à necessidade crescente de sistemas capazes de lidar com a complexidade e o volume de dados em diferentes contextos, desde bibliotecas e arquivos até plataformas digitais e redes colaborativas de pesquisa.

Os SOC são os instrumentos que apoiam a sistematização e organização do conhecimento. Para Carlan (2010), os SOC são instrumentos que traduzem sistematicamente os conteúdos de documentos originais para estruturas organizadas, representando seu conteúdo com o objetivo de facilitar a organização da informação e do conhecimento, resultando em processos mais eficientes de recuperação das informações descritas nos documentos.

A institucionalização social da organização e representação do conhecimento é legitimada assim pelos SOC, que atuam como instrumentos essenciais para viabilizar, em última instância, a recuperação da informação em diversos sistemas de informação. Exemplos de SOC incluem sistemas de classificação, listas de cabeçalhos de assuntos, tesouros, taxonomias e ontologias. Esses instrumentos compartilham características comuns, como a identificação da extensão e da intenção dos conceitos, a organização das relações conceituais e a ênfase no controle terminológico (MOREIRA, 2010).

No aspecto pragmático e estrutural, os SOC baseiam-se na compreensão das relações entre o termo (expressão) e o conceito (conteúdo), bem como nas conexões entre os próprios conceitos (sistemas nocionais), que determinam a identidade representacional de cada termo. Esses conhecimentos são indispensáveis para o desenvolvimento de sistemas de representação avançados (MOREIRA, 2010).

Dessa maneira, alguns sistemas de representação avançados, nos últimos anos, vêm se voltando para as tecnologias digitais, onde ao associar o campo de OC, e mais precisamente os SOC, no universo da Ciência da Informação, ressalta-se a importância dos recentes estudos na área de IA, propiciando uma série de discussões em diversas outras áreas.

Os estudos sobre IA, como campo emergente inicialmente associado à Ciência da Computação, tem avançado significativamente graças ao progresso tecnológico das últimas décadas, resultando no desenvolvimento de diversas tecnologias que utilizam essa área e influenciam diretamente o dia a dia das pessoas (CONEGLIAN *et.al.*, 2023).

Os principais componentes e áreas que fundamentam a estruturação de conceitos e definições, no campo da IA, incluem o processamento de linguagem natural, aprendizado de máquina, aprendizado profundo, sistemas especialistas, robótica, dispositivos inteligentes (como assistentes virtuais Siri, Alexa e Cortana), *big data*, computação em nuvem e a Internet das Coisas, entre outras tecnologias em constante desenvolvimento e aprimoramento (GONTIJO e ARAÚJO, 2021).

Sobre estas considerações, e em linguagem contemporânea, a IA pode ser definida então como um código de *software* que possui a capacidade de processar entradas verbais, identificar palavras-chave, reorganizá-las com base em *scripts* previamente programados e gerar respostas que apresentam aparência de inteligibilidade (GUNKEL, 2017).

As IA relacionam-se ainda “[...] con la capacidad de las máquinas para usar algoritmos, aprender de los datos y utilizar lo aprendido en la toma de decisiones tal y como lo haría un ser humano” (ROUHIAINEN, 2018).

Nesse sentido, e tomando por base o atual contexto de dados e informações nato-digitais ou convertidos em formatos para ambientes digitais (digitalizados), têm apontado inúmeras dificuldades em relação à identificação, acesso e consequente utilização destes dados e informações.

Em se tratando de ambientes digitais e em face de inúmeras preocupações advindas por meio deste, “[...] o tratamento digital além de agregar as funções do tratamento convencional, também se ocupa do desenvolvimento de novos padrões e protocolos para organização da informação [...]”, onde “[...] a identificação automática ou semi-automática de descritores se torna imprescindível em suportes com grandes volumes de informações,

facilitando a indexação e aumentando o relacionamento semântico do documento” (MARTINS, 2010:3).

Quanto ao tratamento digital de informações, incluindo todas as etapas com vistas à sua organização, torna-se necessário esclarecer sobre a estrutura e funcionamento das IA, especialmente quanto a seus mecanismos de intercomunicação ou interoperabilidade entre sistemas e linguagens, permitindo não somente a estruturação de dados, mas sua completa organização, disponibilização e acesso em ambientes digitais para uso, reuso, apropriação e geração de novas pesquisas.

Nessa perspectiva, a compreensão acerca das tecnologias ou, mais precisamente, “tarefas” associadas às IA, incluem: aprendizagem, raciocínio, planejamento, percepção, compreensão de linguagem e robótica, dependendo de alguns métodos de aprendizagem e operação, como os métodos de aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*) (RUSSEL e NORVIG, 2013).

Para Janiesh, Zschechi e Heinrich (2021), o *machine learning* tem por tarefa treinar algoritmos para prever ou tomar decisões com base em dados, focando em habilitar máquinas a aprender com dados passados e revisados, melhorando seu desempenho ao longo do tempo. Já o *deep learning*, figura como um subconjunto do *machine learning*, utilizando redes neurais com muitas camadas (daí “profundo”), sendo particularmente poderoso para lidar com grandes e complexos conjuntos de dados, como imagens, fala e texto.

Algumas “modelagens” de aprendizagem profunda, permitem que um sistema de computador “se treine” usando dados históricos, reconhecendo padrões e promovendo inferências probabilísticas.

Nada disso seria possível, não fosse a mediação destas modalidades de aprendizado por uma base ou sistema de “linguagem”, ou o chamado Processamento de Linguagem Natural (PLN), que para Bulegon e Moro (2010), trata-se de uma estratégia de mediação, criada não somente para extrair as informações de textos, mas também facilitar a entrada e estruturação de dados, promovendo o uso de textos e escritos (antigos e recentes), sob diversos contextos e abordagens, entre elas: a simbólica, com regras linguísticas estruturadas e sem ambiguidades; a estatística, utilizando modelos matemáticos; a conexionista, relacionando o estatístico com representações do conhecimento; e a híbrida, que combina as demais.

O PLN, ou a chamada linguística computacional, figura então como um subcampo da IA, que visa aprender, compreender, reconhecer e produzir conteúdo em linguagem humana, e mais especificamente “[...] possibilita a extração de informações relevantes (como entidades nomeadas, tópicos e sentimentos expressos) a partir do conteúdo dos documentos” (FELIPE, 2023:24).

Tavares, Meira e Amaral (2020) destacam que um dos usos mais frequentes das técnicas de aprendizado de máquina é a aplicação de algoritmos baseados em *machine learning* para prever e classificar novas situações a partir de experiências previamente conhecidas dentro de um mesmo contexto. Essa abordagem exemplifica o funcionamento do aprendizado supervisionado, no qual o modelo aprende a partir de dados rotulados para realizar previsões.

Na etapa subsequente, é essencial garantir condições específicas para que o processo atinja os resultados esperados. Nesse contexto, as técnicas de PLN desempenham um papel crucial, fornecendo suporte à análise e interpretação de dados textuais, fundamentais para o sucesso do modelo.

Por isso, a aprendizagem de máquina não supervisionada, pelo que se pode observar, necessita de um léxico em constante atualização, ou mais precisamente de um PLN bem estruturado, sendo talvez a parte mais importante.

Paes (2004) descreve o unitermo como um método de arquivamento utilizado em arquivos especiais e especializados, no qual, após a análise dos documentos, destacam-se os elementos identificadores. Esses elementos são transcritos em fichas-índice na forma de palavras-chave, quando extraídos diretamente dos documentos analisados, ou como descritores, caso sejam selecionados a partir de *thesauri* técnicos ou vocabulários controlados.

Esse método, ao se basear na extração e organização de termos-chave, pode ser relacionado ao uso contemporâneo de técnicas de PLN, que também visam identificar e categorizar informações textuais de forma automatizada, promovendo a organização e recuperação de informações de maneira eficiente.

A OC e a IA apresentam uma relação de mútuo benefício, marcada pela intersecção de conceitos e aplicações práticas em diversos contextos. A OC busca estruturar, classificar e representar o conhecimento de forma sistemática, com o objetivo de facilitar o acesso e a recuperação de informações.

Por sua vez, a IA, com suas capacidades avançadas de processamento e análise de dados, oferece ferramentas que potencializam os processos envolvidos na OC, contribuindo para uma gestão mais eficaz e dinâmica do conhecimento.

Na OC, sistemas como taxonomias, tesauros, ontologias e outros SOC desempenham um papel essencial na categorização e recuperação da informação. A IA complementa essas funções ao oferecer tecnologias baseadas em aprendizado de máquina e PLN, capazes de lidar com grandes volumes de dados e identificar padrões de maneira automatizada.

O PLN é um exemplo claro dessa colaboração, sendo utilizado para extrair informações relevantes de documentos, categorizar conteúdos e gerar descritores automaticamente, permitindo não apenas o tratamento de dados textuais, mas também a criação de estruturas semânticas que conectam informações de forma significativa.

No futuro, espera-se que essa integração evolua ainda mais, com o desenvolvimento de sistemas híbridos que combinem o raciocínio lógico e semântico da OC com as capacidades preditivas e adaptativas da IA. Essa colaboração promete atender às demandas de um mundo digital em constante expansão, promovendo o acesso equitativo à informação e o avanço do conhecimento em escala global.

4. Organização do Conhecimento e Arquivologia

A OC é um campo interdisciplinar que busca estruturar, classificar e representar informações, facilitando seu acesso, recuperação e uso.

Hjørland (2008) e Dahlberg (2006) oferecem definições centrais para o conceito, destacando a OC como uma prática que abrange desde a catalogação e indexação até a elaboração de sistemas complexos de classificação, com impactos significativos em diversas áreas do conhecimento, incluindo a Arquivologia.

Historicamente, a organização do conhecimento nos arquivos remonta aos primeiros registros escritos, como os arquivos de Ebla, na antiga Mesopotâmia, onde tabletas de argila continham registros econômicos e administrativos organizados em categorias funcionais.

Outro exemplo são os templos egípcios, que preservam documentos relacionados a transações comerciais e rituais religiosos, demonstrando um esforço sistemático para classificar e armazenar informações essenciais para a gestão social e cultural. Esses exemplos ilustram como, desde as primeiras civilizações, os arquivos serviram não apenas como depósitos de documentos, mas também como sistemas de OC destinados a sustentar as estruturas de poder e a memória coletiva.

Com o avanço das práticas arquivísticas na Idade Moderna, consolidaram-se os princípios de respeito à proveniência e à ordem original, que continuam sendo pilares fundamentais na descrição e classificação de documentos.

A incorporação de princípios da OC no âmbito arquivístico foi intensificada no século XX, com a crescente necessidade de sistematizar e tornar acessíveis grandes volumes de documentos. Autores como Paul Otlet, que idealizou o conceito de "Mundaneum" como um repositório universal de conhecimento, influenciaram significativamente as abordagens de organização documental.

Embora os princípios e conceitos fundamentais da Arquivologia – como o conceito de fundo, o princípio da proveniência e o conceito de documento de arquivo – permaneçam relevantes e devam ser rigorosamente observados, a área tem gradualmente se afastado do paradigma custodial. Essa mudança direciona o foco para uma abordagem que, além de valorizar a custódia, enfatiza o caráter social inerente aos processos de criação, organização e uso de documentos arquivísticos (TOGNOLI, SILVA e SILVA, 2019).

A aproximação entre as duas áreas, notadamente por meio do conceito de conhecimento arquivístico, permite que a Arquivologia se beneficie dos estudos já consolidados da área de OC, em especial aqueles relacionados aos dilemas éticos nos processos de organização e representação do conhecimento, à indexação e aos formatos e padrões interoperáveis para representação de dados (TOGNOLI, SILVA e SILVA, 2019).

Paralelamente, iniciativas como as normas padronizadas promovidas pelo Conselho Internacional de Arquivos, incluindo a ISAD(G), impulsionaram a estruturação e a descrição sistemática de documentos, destacando-se como marcos na evolução das práticas arquivísticas (BARROS, BASTOS e SANTOS, 2022).

Essas contribuições evidenciam a interseção entre teoria e prática no esforço de transformar os arquivos em centros de conhecimento acessíveis e interoperáveis. Ferramentas como tabelas de classificação, tesauros e vocabulários controlados tornaram-se essenciais para padronizar a organização documental (TOGNOLI e BARROS, 2015).

A organização física de documentos, tradicionalmente baseada em métodos como a classificação por fundo ou série documental, é ampliada pelas tecnologias digitais, que possibilitam a criação de metadados detalhados e sistemas de informação que facilitam o acesso remoto e a interconexão de dados. Com a digitalização e o surgimento de sistemas eletrônicos de gestão arquivística, os SOC passaram a desempenhar um papel central na estruturação e recuperação da informação.

Conforme a própria trajetória da organização e representação do conhecimento, a maior preocupação é o conteúdo e sua representação – ainda que, no caso dos arquivos, o contexto seja o elemento primordial dessa organização.

Nos arquivos, a organização e a sistematização do conhecimento desempenham papéis essenciais tanto para documentos físicos quanto para os digitais. O contexto de produção de documentos inclui fatores como a proveniência, a finalidade administrativa e o valor histórico, que determinam como esses registros devem ser organizados e preservados (TOGNOLI e BARROS, 2015).

A sistematização de informações em arquivos digitais requer uma abordagem mais abrangente, considerando a complexidade dos dados e as demandas por interoperabilidade entre diferentes sistemas. Para além de garantir a integridade e a autenticidade dos documentos, os arquivistas devem trabalhar na construção de estruturas lógicas que permitam a recuperação contextualizada da informação, respeitando os princípios arquivísticos, como a organicidade e a unicidade (BARROS, BASTOS e SANTOS, 2022).

A descrição é um dos processos fundamentais para a organização e sistematização do conhecimento em arquivos. Nesse contexto, a descrição arquivística passou a ter como objetivo adicional facilitar o acesso aos documentos de maneira mais eficiente. Ao longo dos anos, então, este processo foi sendo readaptado e redirecionado, como destaca Duranti:

[...] a descrição vem perdendo progressivamente seus objetivos de prestar contas dos fundos da instituição arquivística, facilitando a recuperação física de documentos e fornecendo acesso à informação. Esses objetivos são agora preenchidos pela compilação de registros de acesso, listas de localização, índices e todos os tipos de instrumentos de recuperação. O principal objetivo de escrever sobre os documentos é destacar as relações contextuais e de proveniência (DURANTI, 1993:50).

A partir do século XX, a descrição arquivística foi além da representação dos conteúdos dos documentos passando, também, a priorizar a relação dos documentos com seus respectivos contextos. Esse processo começou a ser realizado em paralelo à classificação, com a finalidade de melhor representar o documento, integrando conteúdo (assunto) e contexto (informações de produção). Sem dúvida, a compreensão da descrição e classificação enquanto atividades conjuntas possibilita a existência de relações interdisciplinares entre a Arquivologia e a OC, possibilitando à primeira novas abordagens teórico-metodológicas.

Aliado a isto, Guimarães e Tognoli (2015) tratam a própria organização do conhecimento arquivístico enquanto domínio, pontuando:

É importante observar que a organização do conhecimento arquivístico tem seu núcleo na classificação (estrutural ou funcional) e nos procedimentos de descrição, com o objetivo de estabelecer prerrogativas de conhecimento. Essas

prerrogativas são construídas de acordo com o contexto da criação dos documentos, com o objetivo de garantir seus valores probatórios e históricos (GUIMARÃES e TOGNOLI, 2015:567).

Schellenberg (1956) destaca a importância da descrição arquivística como um processo que organiza e contextualiza os registros, permitindo sua compreensão no âmbito de sua proveniência. Essa prática encontra paralelos na OC, que também busca compreender os contextos de produção e uso da informação para criar sistemas de classificação efetivos.

Além disso, a incorporação de tecnologias digitais em ambos os campos tem promovido a adoção de soluções inovadoras, como ontologias e metadados, que ampliam as possibilidades de organização, estruturação, acessibilidade e preservação dos documentos em arquivos.

Em termos de SOC, observa-se uma interseção entre as práticas arquivísticas e as ferramentas epistemológicas utilizadas para categorizar, indexar e descrever informações em diversas áreas do conhecimento. Alguns instrumentos de OC, como tesouros, taxonomias e ontologias, compartilham com os instrumentos arquivísticos o objetivo de facilitar a navegação e a recuperação de informações. Nesse sentido, a padronização proporcionada pelos vocabulários controlados e pelos planos de classificação é essencial para a interoperabilidade entre sistemas de informação, assegurando que usuários de diferentes contextos possam acessar e interpretar os conteúdos arquivísticos.

Alguns instrumentos como guias, catálogos e índices, desempenham um papel crucial na acessibilidade, servindo não apenas para localizar informações, mas também para contextualizar os documentos dentro de suas funções administrativas e históricas. Dessa forma, eles permitem que o conhecimento contido nos arquivos seja acessível não apenas para especialistas, mas também para públicos mais amplos, atendendo ao princípio democrático de acesso à informação.

Como dito anteriormente, as tecnologias digitais têm ampliado as possibilidades de acesso e recuperação de informações. Ferramentas baseadas em IA, por exemplo, podem automatizar a criação de índices e guias, personalizar a recuperação de informações e até identificar relações entre documentos que não seriam imediatamente visíveis.

No entanto, a integração dessas tecnologias deve ser guiada pelos princípios éticos da Arquivologia e pela garantia de um acesso equitativo à informação. Acredita-se aqui, portanto, em uma OC integradora que busque contribuir para além de seu próprio campo de atuação, visando auxiliar na construção de melhores SOC no âmbito dos arquivos.

5. Inteligência Artificial e possíveis contribuições à Arquivologia

Em relação às IA, o desenvolvimento e incremento crescente de tecnologias e ferramentas tem proporcionado agilidade e rapidez de inúmeras atividades por meio de processos automatizados inteligentes, incidindo de maneira significativa no cotidiano de várias outras ciências, entre elas, a Arquivologia, contribuindo assim, na organização e tratamento da informação arquivística.

Visto por esse lado então, as ferramentas de IA vêm sendo integradas aos sistemas de informação, otimizando processos como a categorização automática de documentos, a extração de metadados e a análise de conteúdo, o que é reforçado por Gomes, Fernández Marcial e Santos (2021) ao destacar que,

[...] o impacto da IA é de primeira ordem nas tarefas relacionadas com o processamento da informação, não só no trabalho que pode ser considerado mais mecânico, [...] mas também para os que exigiam maior intervenção humana, como a indexação, a atribuição de metadados ou a descrição de conteúdo (GOMES, FERNÁNDEZ MARCIAL e SANTOS, 2021:403).

E, complementado por Gomes (2022:12), reiterando que “a classificação de conteúdos com recurso a ontologias, produzidas tanto de forma manual quanto automática, a criação automatizada de metadados ou o Processamento em Linguagem Natural (PNL) são úteis aos arquivos, bibliotecas e museus”.

Nessas perspectivas e no âmbito das tarefas mais mecânicas, como a digitalização e o reconhecimento óptico de caracteres (OCR), a IA trouxe avanços significativos, reduzindo custos e aumentando a eficiência. Essas ferramentas, ao automatizar processos antes realizados manualmente, permitem que os profissionais da informação se concentrem em atividades mais complexas e estratégicas.

Contudo, o impacto mais disruptivo ocorre em atividades que tradicionalmente demandavam maior expertise humana, como a atribuição de metadados, a indexação e a descrição arquivística. Ferramentas baseadas em PLN e aprendizado profundo (*deep learning*) podem gerar metadados automaticamente, analisando o conteúdo textual ou visual dos documentos e sugerindo descrições com base em padrões detectados em grandes bases de dados. Este avanço representa uma nova fase para a Arquivologia e a Ciência da Informação, na qual a IA complementa ou até substitui parcialmente o papel humano na criação de estruturas descritivas.

No tocante ao uso de tecnologias de IA na Arquivologia, algumas soluções ou proposições têm sido idealizadas e utilizadas, especialmente nos aspectos de “[...] transparência de registros e acesso à informação arquivística, assim como a recuperação de grandes quantidades de dados para sua ordenação” (DORNELES, 2022:12).

A IA contribui significativamente para a transparência ao facilitar o acesso rápido e preciso a registros arquivísticos. Ferramentas baseadas em aprendizado de máquina podem ser empregadas para automatizar a classificação de documentos, identificar padrões em grandes volumes de dados e até mesmo gerar metadados a partir de conteúdo não estruturado, como imagens e manuscritos digitalizados.

No entanto, o uso de IA para acesso à informação arquivística não está isento de desafios. É necessário garantir que os algoritmos empregados sejam transparentes em sua lógica e que não reproduzam vieses que possam prejudicar a equidade no acesso à informação.

A IA também é uma aliada no gerenciamento de volumes crescentes de dados arquivísticos. Algoritmos de PNL, por exemplo, podem ser utilizados para analisar e categorizar documentos automaticamente, reduzindo significativamente o tempo e o esforço humano envolvidos nesses processos.

Além disso, técnicas de análise de *big data* permitem identificar conexões entre registros que podem enriquecer a contextualização e o uso da informação. Apesar dos benefícios, a integração da IA nessas tarefas levanta questões éticas e epistemológicas. Como os algoritmos de IA são treinados em grandes conjuntos de dados, há o risco de que eles reproduzam vieses inerentes a essas fontes, resultando em descrições ou indexações enviesadas.

A integração de IA na Arquivologia também levanta questões éticas, especialmente no que tange à privacidade dos indivíduos e à preservação da autenticidade dos documentos. Como os algoritmos podem tomar decisões com base em grandes conjuntos de dados, é crucial garantir que a integridade arquivística não seja comprometida.

Além disso, outras aplicações mais específicas de ferramentas de IA também são utilizadas, especialmente as que “[...] tratam de funções arquivísticas fundamentais do trabalho de gestão documental, tais como processamento de fluxos de documentos até o arquivamento, classificação e avaliação arquivística” (DORNELES, 2022:12).

De modo prático, “[...] o Aprendizado de Máquina (Machine Learning), por exemplo, permite que algoritmos sejam treinados para reconhecer padrões e características específicas em documentos digitais, facilitando a classificação e organização de arquivos” (FELIPE, 2023:24).

A exemplo disso, o uso de ferramentas baseadas em IA, como Grammarly, Google AutoML, IBM Watson Discovery e Microsoft Azure AI, já está transformando os processos de criação, estruturação e classificação de documentos arquivísticos. Essas tecnologias oferecem soluções avançadas que aprimoram tanto a qualidade da produção documental quanto a eficiência na organização e recuperação de informações.

O Grammarly, amplamente utilizado para revisão e edição de textos, também desempenha um papel relevante na criação de documentos arquivísticos. Sua aplicação pode garantir a clareza, a consistência e a padronização textual, características essenciais na gestão documental.

Segundo Boudry e Peeters (2020:29), “[...] a aplicação de ferramentas de escrita assistida como Grammarly reduz inconsistências e melhora a legibilidade, atributos indispensáveis para documentos oficiais e administrativos”.

Essa ferramenta também pode ser configurada para sugerir termos padronizados e evitar ambiguidades, auxiliando na conformidade com normas e diretrizes institucionais. No contexto arquivístico, isso é fundamental para a criação de documentos que estejam alinhados com os padrões estabelecidos, como a norma internacional ISAD(G), voltada para a descrição arquivística.

Ferramentas como Google AutoML, IBM Watson Discovery e Microsoft Azure AI representam avanços significativos na classificação automática de documentos arquivísticos. Essas tecnologias utilizam aprendizado de máquina para analisar grandes volumes de dados e identificar padrões, atribuindo classificações e metadados com base em conteúdo, contexto e estrutura.

O Google AutoML, por exemplo, permite que organizações treinem modelos personalizados para classificação documental, sem a necessidade de conhecimentos

avançados em programação. Isso é especialmente útil em contextos arquivísticos, onde os documentos possuem características específicas que podem não ser atendidas por soluções genéricas.

Conforme descrito por Pal, Tiwari e Kulkarni (2019:47), "[...] o Google AutoML democratiza o acesso à IA ao permitir que especialistas em domínios específicos personalizem soluções para seus contextos únicos".

O IBM Watson Discovery se destaca na extração de *insights* e metadados a partir de textos complexos, como relatórios e registros administrativos. Essa ferramenta não apenas organiza os documentos, mas também identifica relações semânticas entre eles, facilitando a recuperação da informação. Já o Microsoft Azure AI oferece serviços integrados que automatizam tarefas de classificação e indexação, promovendo maior eficiência na gestão de arquivos.

A adoção dessas tecnologias no campo arquivístico apresenta múltiplas vantagens, incluindo a redução de tempo e esforço humano, o aumento da precisão na classificação e a capacidade de lidar com grandes volumes de documentos. Contudo, a integração dessas ferramentas também exige atenção a desafios éticos e operacionais. É essencial garantir a qualidade e a confiabilidade dos dados usados para treinar os algoritmos, a fim de evitar vieses e classificações inadequadas.

Por outro lado, existem na atualidade ferramentas tecnológicas destinadas ao tratamento e recuperação de documentos manuscritos, a exemplo da ferramenta Transkribus, desenvolvida pela Universidade de Innsbruck, na Áustria, e disponibilizada no *site* <https://readcoop.eu/transkribus/>, com a participação e colaboração de várias instituições por meio do Projeto READ.

Uma destas instituições é a Universidade Politécnica de Valência (Espanha), que atua com o objetivo de desenvolver novas ferramentas para transcrição e indexação automática de documentos manuscritos compreendidos entre os séculos XIV e XX de vários países como: Finlândia, Alemanha, Itália, Inglaterra, França, Espanha etc.

Com base na pesquisa realizada pelo projeto *transScriptorium*¹, o projeto READ definirá novos padrões em Reconhecimento de Texto Manuscrito, Identificação de Palavras-Chave, Análise de *Layout*, Identificação Automática de Escritores e campos relacionados. O objetivo principal do READ é fornecer uma plataforma de serviço² para o reconhecimento, transcrição e busca automatizados de documentos históricos. Treinar computadores para ler texto manuscrito dessa forma promete revolucionar o acesso aos registros escritos de nossa história³.

Essa experiência já encontra-se consolidada também em ambientes especializados como o Arquivo Histórico de Amsterdão, onde:

A plataforma em linha permite aos utilizadores formar um modelo de reconhecimento de escrita à mão por computador para transcrever

¹ <http://transcriptorium.eu/>.

² <http://transkribus.eu>.

³ https://eadh.org/projects/read_2024.

documentos históricos escritos à mão numa variedade de línguas europeias. (...) Uma vez formado, o modelo utiliza a aprendizagem automática para comparar os padrões de caligrafia que agora conhece, com os dos documentos que o utilizador quer transcrever. O modelo transcreve automaticamente linha a linha. (...) Depois do Transkribus ter feito o seu trabalho, os utilizadores muitas vezes só precisam rever para corrigir quaisquer erros menores. Embora isto possa parecer muito trabalho inicial, pode poupar a arquivistas, historiadores e estudiosos, centenas - se não milhares - de horas sentados em frente a um computador a transcrever à mão o conjunto completo de documentos” (BURKE, 2020 *apud* GOMES, 2022).

No Brasil, iniciativas de igual teor encontram-se em andamento, já com alguns resultados demonstrados, destacando-se os trabalhos do Grupo de Pesquisas paleográficas, filológicas e históricas, “Memória em papel”, do Programa de Pós-graduação em Língua e Cultura da Universidade Federal da Bahia (UFBA?).

“Memória em Papel” é um grupo multidisciplinar que tem como objetivo principal auxiliar na salvaguarda da memória e da história através do desenvolvimento de pesquisas filológicas, paleográficas e históricas. Nossa mais significativa contribuição se apresenta através do desenvolvimento de projetos de edição de textos manuscritos históricos que envolvem análises sob a perspectiva de diversas outras ciências: Diplomática, Codicologia, História, Linguística, Educação, Humanidades Digitais em convênio com a Universidade de Coimbra e a Universidade Nova de Lisboa (Projeto *Memória em papel*, 2024).

Desde 2019, as pesquisadoras Lúcia Furquim Werneck Xavier, Livia Borges Souza Magalhães e Alícia Duhá Lose estão coordenando a constituição de um modelo de leitura automatizada em língua portuguesa para o aplicativo Transkribus, compondo, mais especificamente, o projeto *Muitas Mãos e Muitas Penas* (MAGALHÃES, 2021). Neste projeto, os pesquisadores têm treinado a ferramenta para reconhecer as particularidades da língua portuguesa, de modo que o sistema seja capaz de realizar, de forma automatizada e ágil, a edição de conjuntos documentais previamente selecionados. O foco do trabalho são os documentos do período inquisitorial disponibilizados pelo Arquivo Nacional da Torre do Tombo, em Portugal (MAGALHÃES, 2021).

As etapas de trabalho são divididas então: em uma primeira etapa, que é a digitalização dos manuscritos, seguida pela etapa de reconhecimento do texto, onde:

A plataforma está programada dentro daquilo que chamamos de Handwritten Text Recognition (HTR). Essa tecnologia usa a base do desenvolvimento do Optical Character Recognition, responsável por tomar imagens digitalizadas de impressos e convertê-las em texto editável, num processo que ocorre porque a máquina compara os caracteres com modelos que estão registrados no banco de dados. O HTR é a construção de um banco de dados de reconhecimento de escrita manuscrita, [...] o que demanda, então, um fazer paleográfico para tratar de todas essas questões e, assim, construir um banco de dados realmente válido (MAGALHÃES, 2021:290).

O trabalho desenvolvido pela equipe do Grupo de Pesquisa “Memória em Papel” tem como objetivo construir um banco de dados que permita ao Transkribus operar com documentos em língua portuguesa. Para isso, as imagens selecionadas precisam ser identificadas, seja

pelo paleógrafo ou automaticamente pela ferramenta, indicando precisamente as áreas onde há texto, denominadas de "regiões de texto".

Em seguida, inicia-se o processo de inserção de dados, que consiste na transcrição fiel das palavras contidas nas regiões identificadas, preservando todos os detalhes e adotando uma abordagem conservadora. Após essa etapa, realiza-se a rotulação do que é denominado *ground truth* (verdade fundamental), uma validação do usuário que confirma para a ferramenta que o fólio está pronto para ser incorporado ao banco de dados (MAGALHÃES, 2021).

Quanto mais fólios forem transcritos, mais robusto será o quantitativo de *ground truth* e, conseqüentemente, mais consistente será o banco de dados para construção da automação de transcrições. É, efetivamente, um processo de treinamento. No projeto *Muitas Mãos e Muitas Penas*, tem-se tomado o percentual de 30% de todo conjunto documental inserido no Transkribus (MAGALHÃES, 2021).

Após a construção do banco de dados, inicia-se o desenvolvimento do modelo de leitura automatizada. A plataforma recomenda que este processo seja realizado apenas quando houver entre 5.000 e 150.000 palavras transcritas como *ground truth*. O treinamento é acessado na opção *Train* das ferramentas, onde formulários coletam informações sobre o modelo a ser criado. Embora o processo seja demorado, devido à articulação de redes neurais que reconhecem padrões e classificam dados de forma contínua para aprimoramento, o resultado final é um modelo pronto para uso eficiente (MAGALHÃES, 2021).

E como contribuição mais ampla em termos de gestão documental, as ferramentas de IA podem auxiliar na automação de tarefas rotineiras e complexas, como a identificação de documentos duplicados, o reconhecimento de padrões e o suporte à tomada de decisões, resultando em maior eficiência e precisão no tratamento de grandes volumes de dados, alinhando-se aos princípios dos SOC para garantir a acessibilidade e a preservação da informação ao longo do tempo.

6. Considerações finais

Acreditamos que a adoção de ferramentas de IA no campo arquivístico oferece um potencial expressivo para transformar as práticas de organização, tratamento, gestão e recuperação da informação. A integração dessas tecnologias aos SOC não apenas traz celeridade aos processos de classificação e busca, mas também possibilita novas formas de interpretar e estruturar o conhecimento arquivístico.

A Ciência da Informação, dessa maneira, desempenha um papel essencial na organização, classificação e recuperação da informação, permitindo a compreensão dos dados e informações que alimentam a IA, garantindo sua qualidade e relevância na construção de novos conhecimentos.

A implementação de algoritmos de *machine learning* e *deep learning* em atividades arquivísticas pode promover avanços em várias frentes: desde a automatização de tarefas repetitivas até o desenvolvimento de esquemas de representação mais precisos e

personalizados, que permitam uma recuperação de dados de modo mais eficaz e adequado às necessidades dos usuários.

Esses avanços apontam para um futuro em que os Arquivos públicos e privados se tornem ainda mais acessíveis, eficientes e democráticos, cumprindo um papel fundamental na preservação e disseminação do conhecimento.

Tudo isso, aliado aos SOC que se beneficiam das abordagens de *machine learning* e *deep learning*, proporcionando maior eficiência na classificação e descrição de documentos, além de possibilitar a geração de novos subprodutos informacionais.

Diante do exposto, conclui-se que a colaboração entre IA e SOC apresenta um cenário promissor para a Arquivologia, ampliando o alcance da Ciência da Informação. Compreender a IA e sua aproximação com a estruturação e organização da informação em arquivos por meio de SOC, permite o desenvolvimento de amplas discussões e contribuições para o aprimoramento dos processos arquivísticos e gestão documental, garantindo maior agilidade e precisão no tratamento e organização documental.

Referências bibliográficas

BARROS, T. H. B., BASTOS, C. M. C.; SANTOS, A. C. R.

2022 Sistemas de Organização do Conhecimento no contexto da Arquivologia: aportes metodológicos para seu desenvolvimento. *Acervo*. [Em linha]. 2022 35:1 (2022) 1-20. [Consult. 20 dez. 2024] Disponível em: <https://revista.an.gov.br/index.php/revistaacervo/article/view/1812>.

BOUDRY, F.; PEETERS, P.

2020 Writing assistance in digital workspaces: Enhancing clarity and consistency. *Journal of Technical Communication*. [Em linha]. 2020. [Consult. 20 dez. 2024]. Disponível em: <https://techniumscience.com/index.php/socialsciences/issue/view/42>.

BUCKLAND, M. K.

1991 Information as thing. *Journal of the American Society for Information Science*. [Em linha]. 1991 42:5 (1991) 351-360. [Consult. 10 dez. 2024]. Disponível em: Disponível em: [https://doi.org/10.1002/\(SICI\)1097-4571\(199106\)42:5](https://doi.org/10.1002/(SICI)1097-4571(199106)42:5).

BULEGON, H.; MORO, C. M. C.

2010 Mineração de texto e o processamento de linguagem natural em sumários de alta hospitalar. *Journal of Health Informatics*. [Em linha]. 2:2 (2010). [Consult. 12 dez. 2024]. Disponível em: <https://jhi.sbis.org.br/index.php/jhi-sbis/article/view/5>.

CARLAN, E.

2010 *Sistemas de Organização do Conhecimento: uma reflexão no contexto da Ciência da Informação*. [Em linha]. Brasília, 2010. [Consult. 10 dez. 2024]. Disponível em: <https://periodicos.unb.br/index.php/RICI/article/view/2049>.
Dissertação de mestrado - Universidade de Brasília.

CONEGLIAN, C. S. [et al.]

2023 Inteligência artificial generativa e recuperação da informação: tendências e oportunidades de pesquisa. In ENANCIB - ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO, 23º, Aracaju, 2023 - *Anais...* [Em linha]. Aracaju: ANCIB, 2023. [Consult. 10 nov. 2024]. Disponível em: <https://conferencias.ancib.org/index.php/enancib/xxxiiienancib/paper/viewFile/1944/1401>.

CRESWELL, J. W.

2007 *Projeto de pesquisa: métodos qualitativo, quantitativo e misto*. Porto Alegre: Artmed, 2007.

DAHLBERG, I.

2006 *Knowledge Organization*. [Em linha]. 2006. [Consult. 1 nov. 2024]. Disponível em: http://www.db.dk/bh/lifeboat_ko/CONCEPTS/knowledge_organization_Dahlberg.htm.

DAHLBERG, I.

1995 Current trends in Knowledge Organization. In *Organización del conocimiento en sistemas de información y documentación*. Org. Francisco Javier García Marco. Zaragoza: Universidad de Zaragoza, 1995, p. 7-26.

DORNELES, S. L.

2022 Aplicações de inteligência artificial na gestão documental: revisão sistemática na lista, Emerald e Scopus. *Revista Fontes Documentais*. [Em linha]. 5:Edição Especial-II ABM (2022). [Consult. 1 jul. 2024]. Disponível em: <https://periodicos.ufba.br/index.php/RFD/article/view/57966/31312>.

DURANTI, L.

1993 Origin and development of the concept of archival description. *Archivaria*. [Em linha]. 35 (Spring 1993) 47-54. [Consult. 10 out. 2024]. Disponível em: <https://archivaria.ca/index.php/archivaria/article/view/11884>.

FELIPE, E. Q. S.

2023 *Uso das plataformas de linguagem natural como ferramentas na indexação de arquivos digitais*. [Em linha]. João Pessoa, 2023. [Consult. 10 out.2024]. Disponível em: <https://repositorio.ufpb.br/jspui/handle/123456789/28344>.

GOMES, L. I. E.

2022 Transformação digital e inteligência artificial nos serviços de informação: inovação e perspectivas para a Ciência da Informação no mundo pós-pandemia. *Revista Ibero-Americana de Ciência da Informação*. [Em linha]. 15:1 (2022) 148-166. [Consult. 9 out. 2024]. Disponível em: <https://doi.org/10.26512/rici.v15.n1.2022.41490>.

GOMES, L. I. E.; FERNÁNDEZ MARCIAL, V.; SANTOS, M. N. M.

2021 O Impacto da inteligência artificial nos serviços de informação: inovação e perspectivas para as bibliotecas. In CONGRESSO ISKO ESPANHA-PORTUGAL, 5º, Lisboa, 2021 - *Organização do Conhecimento no Horizonte 2030: Desenvolvimento sustentável e saúde: Atas...* Ed. Carlos Guardado da Silva, Jorge Revez, Luís Corujo. [Em linha]. Lisboa: Centro de Estudos Clássicos; Colibri, 2021. [Consult. 13 out. 2024]. Disponível em: <https://estudogeral.uc.pt/handle/10316/103253>.

GNOLI, C.

2020 *Integrative Levels Classification (ILC)*. In *ISKO Encyclopedia of Knowledge Organization*. Org. Birger Hjørland, Claudio Gnoli. [S. l.]: ISKO, 2020.

GONTIJO, M. C. A.; ARAÚJO, R. F.

2021 Inteligência artificial: impacto acadêmico e social de sua produção científica. *Ciência da Informação Express*. [Em linha]. 2 (2021) 1-6. [Consult. 10 out. 2024]. Disponível em: <https://cienciainformacaoexpress.ufla.br/index.php/revista/article/view/29>.

GUIMARÃES, J. A. C.; TOGNOLI, N. B.

2015 Provenance as a Domain Analysis approach in archival Knowledge Organization. *Knowledge Organization*. [Em linha]. 42:8 (2015) 562-569.

GUNKEL, D. J.

2017 Comunicação e inteligência artificial: novos desafios e oportunidades para a pesquisa em comunicação. *Galaxia*. [Em linha]. 34 (2017) 5-19. [Consult. 12 out. 2024]. Disponível em: <https://www.scielo.br/j/gal/a/K7cGL4Jgp8dXYwzDPZPy6qS/?lang=pt&format=pdf>.

HJØRLAND, B.

2016 Knowledge organization (KO). In *Encyclopedia of Knowledge Organization*. Org. Birger Hjørland, Claudio Gnoli. [Em linha]. [S. l.]: ISKO, 2016, 475-484. [Consult. 20 out. 2024]. Disponível em: http://www.isko.org/cyclo/knowledge_organization.

HJØRLAND, B.

2008 What is knowledge organization (KO)? *Knowledge Organization*. 35:2/3 (2008) 86-101.

JANIESCH, C.; ZSCHECHI, P.; HEINRICH, K.

2021 Machine learning and deep learning. *Electronic Markets*. [Em linha]. 31 (2021) 685-695. [Consult. 5 out. 2024] Disponível em: <https://doi.org/10.1007/s12525-021-00475-2>.

LÖSCH, S.; RAMBO, C. A.; FERREIRA, J. L.

2023 A Pesquisa exploratória na abordagem qualitativa em educação. *Revista Ibero-Americana de Estudos em Educação*. [Em linha]. 18 (2023) e023141. [Consult. 9 out. 2024]. Disponível em: <https://doi.org/10.21723/riaee.v18i00.17958>.

MAGALHÃES, L. B. S.

2021 Da 1.0 até a 3.0: a jornada da Paleografia no mundo digital. *Labor histórico*. [Em linha] 7:3 (2021) 279-295. [Consult. 20 out. 2024]. Disponível em: <https://revistas.ufrj.br/index.php/lh/article/view/42368/29251>.

MARTINS, A. L.

2010 Potenciais aplicações da Inteligência Artificial na Ciência da Informação. *Informação & Informação*. [Em linha]. 15:1 (2010) 1-16. [Consult. 19 out. 2024]. Disponível em: <https://ojs.uel.br/revistas/uel/index.php/informacao/article/view/3882>.

MORAES, I. S.

2017 *Os Conceitos de sistemas de organização do conhecimento e linguagens documentárias: análise de domínios nos PPGCIs - UNESP e UFMG*. [Em linha]. Marília, 2017. [Consult. 15 out. 2024]. Disponível em: https://bdtd.ibict.br/vufind/Record/UNSP_b5c76d43ac3cbc09958e51afd553caa8. Dissertação de Mestrado - Universidade Estadual Paulista.

MOREIRA, W.

2010 *A Construção de informações documentárias: aportes da linguística documentária, da terminologia e das ontologias*. [Em linha]. São Paulo, 2010. [Consult. 13 dez. 2024]. Disponível em: <http://www.teses.usp.br/teses/disponiveis/27/27151/tde-05072010-174533/>.
Tese de doutorado - Universidade de São Paulo.

OHLY, H. P.

2023 I. Dahlberg. In *Encyclopedia of Knowledge Organization*. Org. Birger Hjørland, Claudio Gnoli. [Em linha]. [S. l.]: ISKO, 2023. [Consult. 17 out. 2024]. Disponível em: https://www.isko.org/cyclo/knowledge_organization.

PAES, M. L.

2004 *Arquivo: teoria e prática*. 3ª ed. Rio de Janeiro: FGV, 2004.

PAL, K.; TIWARI, S.; KULKARNI, V.

2019 Google AutoML: Simplifying Machine Learning for Non-Experts. *International Journal of Advanced Research in Computer Science*. (2019).

PRODANOV, C. C.; FREITAS, E. C.

2013 *Metodologia do trabalho científico: métodos e técnicas de pesquisa e do trabalho acadêmico*. 2ª ed. Novo Hamburgo: Feevale, 2013.

ROUHIAINEN, L.

2018 *Inteligencia artificial: 101 cosas que debes saber hoy sobre nuestro futuro*. [Em linha]. España: Alienta, 2018. [Consult. 20 dez. 2024]. Disponível em: https://proassetspdlcom.cdnstatics2.com/usuarios/libros_contenido/arxius/40/39_307_Inteligencia_artificial.pdf.

RUSSELL, S.; NORVIG, P.

2013 *Inteligência Artificial: uma abordagem moderna*. 3ª ed. Rio de Janeiro: Elsevier, 2013. [Consult. 13 dez. 2024]. Disponível em: [https://www.kufunda.net/publicdocs/Intelig%C3%Aancia%20Artificial%20\(Peter%20Norvig,%20Stuart%20Russell\).pdf](https://www.kufunda.net/publicdocs/Intelig%C3%Aancia%20Artificial%20(Peter%20Norvig,%20Stuart%20Russell).pdf).

SANTOS, R. F.; NEVES, D. A. B.; SOUZA, E. D.

2019 A Organização do Conhecimento como domínio de estudo da Ciência da Informação: uma reflexão a partir dos aspectos epistemológicos. In *Organização do conhecimento responsável: promovendo sociedades democráticas e inclusivas*. Org. Thiago Henrique Bragato Barros, Natália Bolfarini Tognoli. [Em linha]. Brasília: EDUFPA, 2019, p. 95-104. [Consult. 20 dez. 2024]. Disponível em: <https://lume.ufrgs.br/bitstream/handle/10183/200575/001102758.pdf?sequence=1&isAllowed=y>.

TAVARES, L. A.; MEIRA, M. C.; AMARAL, S. F.

2020 Inteligência artificial na educação: Survey. *Brazilian Journal of Development*. [Em linha]. 6:7 (2020) 48699-48714. [Consult. 15 out. 2024]. Disponível em: <https://doi.org/10.34117/bjdv6n7-496>.

TOGNOLI, N. B.; BARROS, T. H. B.

2015 Os Processos de representação do conhecimento arquivístico: elementos históricos e conceituais da classificação e descrição. In *Organização do conhecimento e diversidade cultural*. Org. José Augusto Chaves Guimarães, Vera Dodebei. [Em linha]. Marília: ISKO-Brasil; FUNDEPE, 2015, p. 94-99. [Consult. 15 out. 2024]. Disponível em: <https://livroaberto.ufpa.br/jspui/handle/prefix/1202>.

TOGNOLI, N. B.; SILVA, A. M. S.; SILVA, S. A. P.

2019 Organização do Conhecimento e Arquivologia: uma análise de domínio nos periódicos Knowledge Organization e Scire. *Informação & Informação*. [Em linha]. 24:3 (2019) 52-77. [Consult. 18 out. 2024]. Disponível em: <https://doi.org/10.5433/1981-8920.2019v24n3p52>.

Jayron Viana dos Santos | jayron.viana@unesp.br

Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Brasil

Sônia Maria Troitiño Rodriguez | sonia.troitino@unesp.br

Universidade Estadual Paulista “Júlio de Mesquita Filho” (UNESP), Brasil