


Artigo

# Engenharia do Conhecimento e Ciência da Informação aplicadas a Objetos Digitais Imagéticos na Era de Dados

Ismaelly Batista dos Santos Silva<sup>1</sup> 

Hildenise Ferreira Novo<sup>2</sup> 

**Resumo:** O papel da Engenharia do Conhecimento, como área científica multidisciplinar, à luz de teorias e práticas em mineração e curadoria de dados digitais, Organização do Conhecimento, programação computacional, desenvolvimento de sistemas e Inteligência Artificial. A Engenharia do Conhecimento que se estabelece socialmente, de modo aplicado ante às contribuições, para fomentar a Era de Dados. A pesquisa em tela objetiva analisar os Objetos Digitais Imagéticos como parte do movimento de Dataficação no contexto de Big Data, além de caracterizar um repertório de processos em Engenharia do Conhecimento: sobretudo àqueles que dialogam fronteiriçamente com saberes e fazeres em Ciência da Informação. Para atender a este objetivo optou-se por uma metodologia ensaística, reunindo fontes de informação sistematicamente levantadas e analisadas mediante os procedimentos científicos de análise de conteúdo. A discussão aponta que as mídias digitais atuam de modo propulsor à proliferação de Objetos Digitais Imagéticos, corroborando com o fenômeno de Big Data; um cenário no qual a Organização do Conhecimento (enquanto processo, na Engenharia do Conhecimento, subsidiado por preceitos comuns à Ciência da Informação) opera para metodizar o encadeamento e viabilizar a recuperação e difusão de dado digitais.

**Palavras-chave:** Organização do Conhecimento; Era de Dados; Objetos Digitais Imagéticos; Engenharia do Conhecimento; Ciência da Informação.

*Knowledge Engineering and Information Science applied to Imagetive Digital Objects in the Data Age*

**Abstract:** The role of Knowledge Engineering, as a multidisciplinary scientific area, in light of theories and practices in data mining and curation, Knowledge Organization, computer programming, system development, and Artificial Intelligence. Knowledge Engineering establishes itself socially, in an applied manner, in response to contributions, to foster the Data Age. The research at hand aims to analyze Imagetive Digital Objects as part of the Datafication movement in the context of Big Data, as well as to characterize a repertoire of processes in Knowledge Engineering: especially those that border on knowledge and practices in Information Science. To meet this objective, an essayistic methodology was chosen, systematically gathering and analyzing sources of information through

<sup>1</sup> Doutora em Ciência da Informação, Universidade Federal da Bahia, ismaellybatista@gmail.com

<sup>2</sup> Doutora em Difusão do Conhecimento, Universidade Federal da Bahia, denisenovo@gmail.com

DOI: <https://doi.org/10.35699/2237-6658.2024.51378>

Revista Múltiplos Olhares em Ciência da Informação, Belo Horizonte, v. 14, e51378, 2024



scientific procedures of content analysis. The discussion suggests that digital media act as a propellant for the proliferation of Imagetic Digital Objects, corroborating with the phenomenon of Big Data; a scenario in which Knowledge Organization (as a process in Knowledge Engineering, supported by principles common to Information Science) operates to methodize the chaining and enable the retrieval and dissemination of digital data.

**Keywords:** Knowledge Organization; Data Age; Imaging Digital Object; Knowledge Engineering; Information Science.

### *Ingeniería del Conocimiento y Ciencia de la Información aplicadas a Objetos Digitales Imágénicos en la Era de los Datos*

**Resumen:** papel de la Ingeniería del Conocimiento, como área científica multidisciplinaria, a la luz de teorías y prácticas en minería y curación de datos digitales, Organización del Conocimiento, programación computacional, desarrollo de sistemas e Inteligencia Artificial. La Ingeniería del Conocimiento se establece socialmente, de manera aplicada ante las contribuciones, para fomentar la Era de los Datos. La investigación en cuestión tiene como objetivo analizar los Objetos Digitales Imágénicos como parte del movimiento de Dataficação en el contexto de Big Data, además de caracterizar un repertorio de procesos en Ingeniería del Conocimiento: especialmente aquellos que dialogan fronterizamente con saberes y prácticas en Ciencia de la Información. Para cumplir con este objetivo, se optó por una metodología ensayística, reuniendo fuentes de información sistemáticamente recopiladas y analizadas mediante procedimientos científicos de análisis de contenido. La discusión señala que los medios digitales actúan como propulsores para la proliferación de Objetos Digitales Imágénicos, corroborando con el fenómeno de Big Data; un escenario en el cual la Organización del Conocimiento (como proceso, en la Ingeniería del Conocimiento, subsidiada por preceptos comunes a la Ciencia de la Información) opera para metodizar el encadenamiento y viabilizar la recuperación y difusión de datos digitales.

**Palabras-clave:** Organización del Conocimiento; Era de los Datos; Imágenes como Objetos Digitales; Ingeniería del Conocimiento; Ciencias de la Información.

**Como citar este artigo:** SILVA, Ismaelly Batista dos Santos; NOVO, Hildenise Ferreira. Engenharia do Conhecimento e Ciência da Informação aplicadas a Objetos Digitais Imagéticos na Era de Dados. **Múltiplos Olhares em Ciência da Informação**, Belo Horizonte, v. 14, p. 1-19, 2024. DOI: 10.35699/2237-6658.2024.51378.

## 1 Introdução

O papel da Engenharia do Conhecimento se estabelece socialmente ante às contribuições para fomentar a Era de Dados. Isto acontece quando levamos em conta a Engenharia do Conhecimento, enquanto uma área científica multidisciplinar<sup>3</sup>, à luz de teorias

---

<sup>3</sup> Optou-se por adotar o termo multidisciplinar e não interdisciplinar, conforme àquilo que é referenciado por autores como Van Harmelen e Fensel (1995), Witten *et al.* (2011) e Schreiber *et al.* (2000). Epistemologicamente, acredita-se que os produtos científicos e tecnológicos demarcados no contexto da área de Engenharia do Conhecimento e na Ciência da Informação partem de uma percepção de diálogo com outras disciplinas, em que a interdisciplinaridade está alinhada à perspectiva de Hilton Japiassu. Em “Interdisciplinaridade e Patologia do Saber”, lançado em 1976, Japiassu posta que, ante uma intersecção, não é possível de se demarcar uma abordagem heterogênea ou partitiva que possa rotular um domínio de ações de um campo, sendo assim adequado se referir à multidisciplinaridade mesmo que tal abordagem não coaduna com autores da Ciência da Informação adotados neste estudo: Borko (1968), Saracevic (1996), Capurro (2003) e Araújo (2018). Contudo, acredita-se que, em termos comparados, a Engenharia do Conhecimento esteja mais próxima da prerrogativa interdisciplinar que a Ciência da Informação. Esta, porém, não é uma discussão que se pretende esvaziar neste momento.

e práticas em mineração e curadoria de dados digitais, Organização do Conhecimento, programação computacional, desenvolvimento de sistemas e Inteligência Artificial de modo aplicado.

A Era de Dados se inscreve de modo social, no movimento de Dataficação, como produto da alta predominância e densidade de dados, sendo demarcada no contexto do fenômeno de *Big Data*. Pode, assim, ser compreendida como um processo derivado do avanço tecnológico, sob a égide das tecnologias aliadas às telecomunicações orientadas à produção, transmissão, recepção, interações e usos de objetos eminentemente em ambiente digital: sejam eles criados ou sejam incorporados ao contexto basilar de sistemas algorítmicos e dispositivos computacionais.

Da produção massiva e volume expressivo de dados digitais em circulação e expansão no *Big Data*, possíveis e necessários à Era de Dados, emergiu o aprimoramento de Objetos Digitais, na modalidade de algoritmos como aplicações, a partir de linguagem de programação computacional voltada à realização de tarefas cada vez mais complexas e retroalimentadas por mecanismos de aprendizagem de máquina no contexto da Inteligência Artificial, na qual os algoritmos coletam, processam e promovem, via declarações lógicas, soluções para demandas coerentes com as necessidades manifestadas em comandos: digitados, voz, clique ou apresentação de imagem no âmbito de um sistema considerado inteligente.

No ambiente das inteligências artificiais, como ferramentas para resolução de tarefas cada vez mais diversificadas, coexiste o processo interativo de retroalimentação de sistemas algorítmicos. Dadas as modalidades de aprendizagem não supervisionada e de aprendizado por reforço, um sistema inteligente pode lançar mão de artefatos disponíveis em seu banco de dados para ampliar sistematicamente sua capacidade lógica de resolver tarefas e ampliar, de modo significativo, as declarações iniciais de sua programação mediante sua função (Witten, *et al.* 2011).

Mecanismos de Inteligência Artificial podem realizar retroalimentação por meio de funções de tentativa e erro, a partir da base de dados disponível de modo local ou remoto, sendo que esse último é possível por interoperabilidade a bancos de dados os quais possa se conectar. Em meio à Era de Dados, o aperfeiçoamento automatizado dos sistemas computacionais considerados inteligentes possui, enquanto insumo particular, o próprio produto de suas interações na resolução de tarefas. Ou seja, tratando-se de usos com objetivo de satisfazer necessidades de seres humanos, quanto mais um sistema é utilizado mais dados são gerados e se tornam acessíveis à máquina como portfólio de respostas “certas” ou

“erradas”, num processo tecnológico em que a máquina aprenderá arranjando as redes neurais digitais presentes em seu código de programação.

Na esteira do desenvolvimento tecnológico e dos mecanismos de telecomunicações figura o papel propulsor das mídias digitais. Elas passaram a compor o fenômeno social já proeminente no estopim dos anos dois mil e denominado como cibercultura (Lévy, 1999). A cultura de socialização, em um ambiente digital e ferramenta para o comércio eletrônico, se tornou acentuada com a escalada de aplicativos e plataformas de serviços como, por exemplo, o *Instagram*, *TikTok* e *Pinterest*.

As mídias sociais digitais operam por meio de aplicativos executáveis em dispositivos móveis, como *smartphones*, e chamam a atenção por explorarem a conectividade e a interatividade em longos períodos. Isso leva ao consumo de dados, fornecidos por suas plataformas e agências, e ao uso e interatividade, os quais implicam na produção de novos dados como fruto da experiência particular de cada usuário, reverberando em milhares de dados coletados e armazenados pelos provedores de internet e bases de dados das empresas responsáveis pelas mídias: sobretudo, adotando o uso de imagens estáticas ou em movimento como principais mecanismos de comunicação, seja nas modalidades de fotografias, ilustrações, memes ou vídeos (Innarelli, 2023).

Há, então, um fenômeno: o protagonismo das imagens em mídias sociais digitais. As imagens passam a operar como Objetos Digitais e têm reafirmado, neste prisma, a comunicação visual em seu apelo dinâmico e referencial marcante, o que em longa data já é adotado por agências de publicidade, telecomunicação e manifestação sociocultural. A possibilidade, agilidade e eficácia da comunicação visual nos cenários computacionais, em meio à Era dos Algoritmos e uso da Inteligência Artificial, abarca possibilidades, mas também desafios, diante da perspectiva de dar ordem a verdadeiros oceanos de dados, que diariamente são produzidos.

O teor dos dados é massivamente produzido entre humanos, com intermédio de máquinas, entre humanos e máquinas ou ainda máquinas e máquinas. Sobretudo no recorte dos Objetos Digitais, como as imagens, se requer que o conhecimento manifestado em dígitos binários seja estruturado, viabilizando, assim, o acesso e recuperação de seu teor informativo. Conforme analisaremos a seguir, a Engenharia do Conhecimento tem demarcado seu diálogo multidisciplinar e operado com relevância, agrupando saberes e práticas de campos cruciais na empreitada da Era de Dados.

A pesquisa em tela configura-se como um estudo inédito, relacionado a um trabalho de doutorado. Apresenta relevância teórica justificada, por se tratar de um tema atual que

requer o estabelecimento de bases epistemológicas. Representa, ainda, de modo significativo, uma emergência social diante dos processos de organização e acesso a Objetos Digitais Imagéticos, que implicam em valioso produto do conhecimento humano na Era de Dados.

Enquanto objetivos foram analisados os Objetos Digitais Imagéticos ante o movimento de Dataficação no contexto de *Big Data*, assim como foi caracterizado um repertório de processos em Engenharia do Conhecimento: sobretudo, àqueles que dialogam, fronteiriçamente, com saberes e fazeres em Ciência da Informação.

Metodologicamente, o estudo é caracterizado como dedutivo (Marques, 1999), sendo tipificado a partir de seus objetivos como descritivo (Gil, 2008). Como procedimento para atender aos objetivos foi adotado um modelo ensaístico que reúne fontes de informação, livros e artigos, considerados seminais à composição de referencial teórico e que coadunam, cientificamente, com os apontamentos apresentados como resultados ao tema proposto.

As fontes de informação fixadas em livros e artigos acadêmicos foram levantadas por meio de busca ativa, em publicações nacionais e internacionais indexadas nas plataformas Scientific Electronic Library Online (SciELO), Academia.edu e na Base de Dados em Ciência da Informação (BRAPCI), identificadas a partir da adoção dos descritores “Engenharia do Conhecimento”, “*Big Data*”, “Era de Dados” “Inteligência Artificial”, “Organização do Conhecimento” e “Objetos Digitais Imagéticos”. O recorte temporal foi fixado pelo critério de pertinência da publicação, mediante o alinhamento do seu conteúdo ao presente estudo.

O produto analítico seguiu a técnica de referência para análise de conteúdo (Bardin, 2011), partindo do levantamento preliminar dos textos, posterior identificação do teor temático e ordem de relevância baseada em número de citações, leitura e esquematização das ideias centrais e inter cruzamento dedutivo dos pontos de contato, sob uma abordagem qualitativa (Minayo; Sanches, 1993). Esta ação possibilitou aos pesquisados, de maneira desafiadora, desde a realização da análise acerca das temáticas centrais até a caracterização das aproximações entre Engenharia do Conhecimento e Ciência da Informação na perspectiva dos Objetos Digitais Imagéticos.

Nas seções seguintes encontra-se um breve aporte teórico, o qual estabelece a atuação da Engenharia do Conhecimento e da Ciência da Informação nos processos de desenvolvimento social e na transformação tecnológica na Era de Dados, onde emergem os Objetos Digitais Imagéticos enquanto produtos do conhecimento em ambiente digital que requer organização para recuperação.

Posteriormente, são apresentados e discutidos os resultados que demarcam aspectos fundamentais dos campos de Engenharia do Conhecimento e Ciência da Informação como

áreas distintas. Aqui se caracteriza a Organização do Conhecimento, sobretudo a partir das Ontologias aplicadas à Inteligência Artificial, revelando o processo fronteiro entre as referidas ciências para organização e recuperação de Objetos Digitais Imagéticos. Por fim, as considerações finais sintetizam a discussão, identificando as limitações e as possibilidades investigativas futuras.

Na sequência são apresentadas as referências que foram utilizadas na feitura desta pesquisa, permitindo a construção do presente texto.

## **2 Referencial teórico**

As ferramentas eletrônicas aliadas à comunicação se intensificaram e passaram a ocupar relevância estratégica na sociedade mundial após a Segunda Guerra, em especial as tecnologias computacionais. Foram elas que, significativamente, contribuíram na transformação das culturas, com saltos críticos de inovação e em um curto espaço de tempo. Também auxiliaram no surgimento de disciplinas científicas despertadas mediante esforços colaborativos de profissionais advindos de áreas já consolidadas, coadunando técnica e teoricamente para o incremento de sistemas eletrônicos orientados à transmissão, armazenamento e recuperação da informação. Foi no caso da Ciência da Informação, na visão de Saracevic (1996), e da Engenharia do Conhecimento, segundo Van Harmelen e Fensel (1995).

O conhecimento se moldou ao longo do tempo, por conta das transformações ocorridas em ambientes de dados computacionais e os serviços de informações em todo mundo, se tornando algo estratégico. Por esta razão, se fez necessário o registro de seus processos técnicos, que asseguram o acesso sistemático ao volume crescente de dados. Uma das frentes tecnológicas que surgiu foi a Engenharia do Conhecimento, cuja “[...] primeira era da tecnologia se estendeu do final dos anos 70 aos meados dos anos 80. Este período foi caracterizado pelo desenvolvimento de novas técnicas de programação” (Van Harmelen; Fensel, 1995, p. 02, tradução nossa).

Tecnologias e insumos computacionais foram se consolidando, empenhados no aprimoramento dos sistemas de dados e no uso desses ambientes para captura e compartilhamento de conhecimento humano. Assim, as necessidades de deslocar o patamar técnico-científico ligado à programação, em especial às linguagens adotadas, e a visão autônoma dos desenvolvedores ligadas às questões que versaram sobre representação semântica, lógica e padrões de codificação foram os fatores que demarcaram os desafios no âmbito da Engenharia do Conhecimento.

Nas primeiras décadas do século XX, o trabalho da Engenharia do Conhecimento foi citado no catálogo de linguagens formais categorizados em três grupos: “[...] (i) linguagens formais baseadas no modelo conceitual KADS, (ii) linguagens formais baseadas em outros modelos conceituais e (iii) o uso de linguagens formais da Engenharia de Software” (Van Harmelen; Fensel, 1995, p. 10, tradução nossa).

A estratégia de desenvolvimento e a implementação de ferramentas foram baseadas no campo da Engenharia do Conhecimento, como soluções às demandas por inovação e organização dos aspectos do conhecimento em ambiente computacional: especificamente, nas crescentes bases de dados e mediante as carências iminentes. Isso se concretizou na diversificação das frentes de atuação e no diálogo com diferentes disciplinas, como a Inteligência Artificial, a Engenharia de *Software*, *Data Mining*, Lógica e Organização do Conhecimento (Witten, *et al.* 2011).

Neste sentido, o modelo conceitual *CommonKADS*, desenvolvido pelo consórcio europeu *Knowledge Acquisition and Design Structuring* (KADS), torna ampla a frente de avanços e partilhamento de metodologias para incremento de soluções na ordem de padronização dos processos de captura, transferência e, principalmente, na modelagem de domínios para constituir especificações de tarefas, em que a representação semântica para compor de forma relevante a linguagem natural passível de ser automatizada e necessariamente inteligível à sistemas computacionais, passam a atuar de modo inteligente (Witten, *et al.* 2011).

O desafio de representação semântica e tradução da linguagem natural via automação para processamento e recuperação de dados em sistemas de agentes inteligentes seguiu a empreitada da Engenharia do Conhecimento nas décadas de 70, 80 e 90 (Van Harmelen; Fensel, 1995).

No primeiro período, Aquisição de Conhecimento foi visto como uma transferência direta de experiência humana em resolução de problemas para um programa de computador. O conhecimento adquirido foi imediatamente representado por um protótipo em execução. No segundo período, a Aquisição de Conhecimento é vista não como um processo de transferência, mas como uma atividade de modelagem. O resultado da aquisição de conhecimento não é mais apenas um programa em execução, mas um modelo conceitual que descreve a experiência em resolução de problemas de maneira independente da implementação. Como consequência, surgiu uma grande lacuna entre o resultado da aquisição de conhecimento e a implementação final de um SBC. Os modelos conceituais descritos em linguagem natural são informações insuficientes para a etapa de implementação, porque são principalmente uma questão de interpretação da linguagem natural se uma implementação atende à tal especificação. A intuição de um programador tem que preencher as lacunas e resolver a

ambiguidade em tais modelos. Normalmente não está claro se um programador possui o domínio e o conhecimento de tarefas necessárias para fazer isso corretamente. O desenvolvimento de técnicas formais de modelagem foi uma resposta natural à esta lacuna, ao definir um nível intermediário. (Van Harmelen; Fensel, 1995, p. 03, tradução nossa)

O desafio de estruturação de redes semânticas aliada à captura do conhecimento na Gestão do Conhecimento e modelagem de domínios especializados à mercê do formalismo necessário para atribuição de lógica às tarefas descritas no contexto da Inteligência Artificial. Esse é o “caminho” que demarca o trabalho com facetas da Organização do Conhecimento, sob as quais as Ontologias figuram de modo imprescindível, pois traduzem, de maneira descritiva, relações e restrições entre dados ou objetos em domínio como sistemas computacionais.

Diferente é a abordagem em Ciência da Informação, em que as Ontologias de Domínio são preteridas ao “[...] especificarem objetos em uma área canônica ou campo do saber como áreas científicas e contextos com potencial abrangência de relações no mesmo domínio” (Silva, 2023, p. 40). Em computação, base dialógica multidisciplinar na qual se inscreve a Engenharia do Conhecimento, as Ontologias típicas e usuais são as que determinam tarefas ou atuam como referência à uma linguagem, como é o caso das Ontologias de Fundamentação.

Faz-se oportuno destacar que, desde o seu surgimento, a Ciência da Informação busca intermediar teorias e práticas para o domínio de sistemas que têm por base a informação em suas múltiplas apresentações, assim como prevê o intercâmbio tecnológico com o campo da Computação no contexto de disciplinaridade (Borko, 1968; Saracevic, 1996; Capurro, 2003).

O eixo investigativo de Organização do Conhecimento na Ciência da Informação encontra-se fixado como o mais remanescente de suas bases de formação, dividindo núcleo epistemológico com a área de Documentação ou Ciências Documentárias e, mais particularmente, a Biblioteconomia em sua abordagem de Classificação e Representação ante à Organização e Tratamento da Informação (Shera, 1980). Aqui, a Organização do Conhecimento tem despontado por sua aplicabilidade e interesse revisitado pelo potencial de contribuição ante às tecnologias da informação Araújo (2018).

Há, geralmente, uma flexibilização no desenvolvimento de sistemas inteligentes para atender especificidades comerciais, como o intercâmbio “[...] entre modelos semiformais e implementações. Eles permitem uma especificação precisa do Sistema Baseado em Conhecimento (SBC) e do conhecimento necessário, ao mesmo tempo, em que abstraem de maneira dupla a implementação” (Van Harmelen; Fensel, 1995, p. 03, tradução nossa).



Tal processo assegura que diferentes níveis de representação semântica possam ser adotados, viabilizando esquemas conceituais robustos e fundamentados como modelo de referência, a exemplo dos geralmente propostos no campo da Ciência da Informação, ao mesmo tempo em que oportuniza que diferentes objetos no domínio dos Objetos Digitais possam ser referenciados pelo modelador ou desenvolvedor do sistema inteligente, pois:

Reflexo do interesse como fenômeno científico e social para diferentes áreas do conhecimento, os Objetos Digitais são identificados e referenciados na literatura científica por meio de termos que, a princípio, parecem equivalentes, sendo alguns mais comuns, como: Artefato Digital, Objeto de Informação e Artefato Tecnológico. [...] Objeto Digital é um Documento Digital, uma vez que possui estrutura algorítmica e o registro manifestado a partir do conhecimento humano (Silva, 2023, p. 29).

Na esteira do desenvolvimento de sistemas, agentes inteligentes e uso de imagens em mecanismos de obtenção de conhecimento e processamento de dados, são os agentes inteligentes desenvolvidos e adaptados para atuar na mineração de dados. Nesse sentido, os componentes de imagens em meio digital, como padrões de pixel e cor, recebem notória atenção no mercado de Inteligência Artificial. As aplicações das ferramentas vão das mídias sociais aos programas estratégicos de Governo (como imagens capturadas por satélites), conforme exemplifica Witten, *et al.* (2011). Ao se reportar ao caso de imagens de triagem o autor revela:

Desde os primórdios da tecnologia de satélite, os cientistas ambientais têm tentado detectar manchas de petróleo a partir de imagens de satélite para alertar antecipadamente sobre desastres ecológicos e impedir o despejo ilegal. Os satélites radar proporcionam uma oportunidade de monitorizar as águas costeiras dia e noite, independentemente das condições meteorológicas. Manchas de óleo aparecem como regiões escuras na imagem, cujo tamanho e forma evoluem dependendo das condições climáticas e do mar. No entanto, outras regiões escuras semelhantes podem ser causadas por condições climáticas locais, como ventos fortes. A detecção de manchas de óleo é um processo manual caro que exige pessoal altamente treinado para avaliar cada região da imagem. Um sistema de detecção de perigos foi desenvolvido para filtrar imagens para posterior processamento manual. Destinado a ser comercializado mundialmente para uma ampla variedade de usuários, agências governamentais e empresas com diferentes objetivos, aplicações e áreas geográficas, este sistema precisa ser altamente personalizável para circunstâncias individuais (Witten, *et al.* 2011, p. 23, tradução nossa).

Seguindo ainda o raciocínio do pesquisador, isso ocorre visto que:

O aprendizado de máquina permite que o sistema seja treinado com base em exemplos de derramamentos e não derramamentos fornecidos pelo usuário e permite que o usuário controle a compensação entre derramamentos não detectados e alarmes falsos. Ao contrário de outras aplicações de aprendizado de máquina, que geram um classificador que é então implantado em campo, aqui é o próprio esquema de aprendizado que será implantado. A entrada é um conjunto de imagens brutas de pixels de um satélite de radar, e a saída é um conjunto muito menor de imagens com supostas manchas de óleo marcadas por uma borda colorida. Primeiro, são aplicadas operações padrão de processamento de imagem para normalizar a imagem. Em seguida, são identificadas regiões escuras suspeitas. Várias dezenas de atributos são extraídos de cada região, caracterizando seu tamanho, forma, área, intensidade, nitidez e irregularidade dos limites, proximidade com outras regiões e informações sobre o fundo nas proximidades da região. Finalmente, técnicas de aprendizagem padrão são aplicadas aos vetores de atributos resultantes (Witten, *et al.* 2011, p. 23, tradução nossa).

Imagens digitais e, por conseguinte, Objetos Digitais na modalidade de documentos de imagens, representam um recurso estratégico de captura do conhecimento (que requer tratamento para viabilizar sua organização para recuperação eficiente), ora têm operado como parte das ações multidisciplinares da Engenharia do Conhecimento: em especial, no cenário de Inteligência Artificial.

Dadas as bases dogmáticas da Ciência da Informação no âmbito dos padrões formais de Sistemas de Organização do Conhecimento se apresenta como um objeto de investigação a ser pensado, principalmente diante das necessidades e oportunidades do mercado e das demandas sociais.

Segundo Innarelli (2023, p.32), “[...] fotógrafos, até então acostumados a preservar seus registros fotográficos em meios analógicos (plástico e papel), foram vítimas da tecnologia digital confiando seus arquivos de registro fotográfico a tecnologias digitais [...]”. Sem conhecimento técnico e infraestrutura adequada, o legado de imagens digitais transformou a agilidade e comodidade em uma realidade problemática no que diz respeito à gestão, organização e preservação da modalidade de documento digital, como é caso das imagens.

### **3 Apresentação e discussão dos resultados**

Esta seção da pesquisa segue o objetivo traçado: analisar o conteúdo descrito nos estudos realizados por autores seminais e contemporâneos do repertório em Engenharia do Conhecimento, que dialogam com a Ciência da Informação. Na perspectiva da análise de conteúdo, sistematizando os vestígios dos assuntos nos documentos (textos), o propósito é se

chegar à interpretação do que foi dito e exposto. Um processo, como já dito por Bardin, dedutivo ou inferencial.

Esta abordagem tem por finalidade efectuar deduções lógicas e justificadas, referentes à origem das mensagens tomadas em consideração (o emissor e o seu contexto, ou, eventualmente, os efeitos dessas mensagens). O analista possui à sua disposição (ou cria) todo um jogo de operações analíticas, mais ou menos adaptadas à natureza do material e à questão que procura resolver (Bardin, 2011, p. 44).

Foram analisados, portanto, os Objetos Digitais Imagéticos como parte do movimento de Dataficação no contexto de *Big Data*, assim como foi caracterizado um repertório de processos em Engenharia do Conhecimento, sobretudo os que dialogam fronteiriçamente com saberes e fazeres em Ciência da Informação.

Procedidas às análises dos conteúdos acerca da Engenharia do Conhecimento sob autoria de Van Harmelen e Fensel (1995), Witten, *et al.* (2011) e Schreiber, *et al.* (2000) está posto e evidenciado seu processo de formação como campo multidisciplinar de abordagem centrada no conhecimento como objeto de estudo, o qual se encontra manifestado em sistemas computacionais, sendo a Engenharia do Conhecimento fronteira e apropriada dos princípios e processos majoritariamente automatizados, a fim de possibilitar metodologias e ferramentas para solução de paradigmas, envolvendo captura, gestão e organização do conhecimento a partir do *Locus* das ciências da computação e em ambiente digital.

Em mesma base analítica, autores como Borko (1968), Saracevic (1996), Capurro (2003) e Araújo (2018) admitem cientificamente que a Ciência da Informação emergiu como uma nova ciência no período do pós-guerra, na década de 1960, pretensa como campo interdisciplinar e cujo objeto de estudo é situado na informação, ao passo que reuniu diferentes abordagens e conceitos de áreas (Matemática, Computação, Sociologia, Psicologia), e, sobretudo, os aspectos epistemológicos das Ciências Documentárias na empreitada da transformação digital através da automação de sistemas de informação. O foco é na infraestrutura para Organização e Tratamento da Informação, visando a recuperação de registros informacionais e considerando aspectos advindos dos fenômenos sociais.

Os objetos de estudo e os percursos de surgimento e desenvolvimento das áreas de Engenharia do Conhecimento e Ciência da Informação demarcam perspectivas distintas, porém fronteiriças, preconizando pontos de contato que mais as aproximam do que repelem. Para ambas, as perspectivas de gêneses estão situadas no advento das Tecnologias da Informação e Comunicação, mais especificamente a popularização dos sistemas computacionais, sendo a Ciência da Informação mais proeminente em seu discurso científico

na década de sessenta (Araújo, 2018), enquanto a Engenharia do Conhecimento desponta por meio dos incrementos possibilitados em Inteligência Artificial na década de oitenta, com a *web* semântica (Van Harmelen; Fensel, 1995; Witten, *et al.* 2011).

Em um prognóstico analítico dedutivo é evidenciado que a Engenharia do Conhecimento avançou, ao longo das décadas, nos cenários de Mineração de Dados, enquanto a Inteligência Artificial no eixo de aprendizagem de máquina, na Representação do Conhecimento e processamento de linguagem natural, lógica e programação computacional, o que leva à infraestrutura, incremento e inovação de sistemas e meios de comunicação. Assim, a Ciência da Informação centrou, em seu desenvolvimento epistemológico diversificado, nos eixos de investigação, alinhado seu percurso de saberes e fazeres na interlocução da Gestão e Organização da Informação e do Conhecimento e utilizando da prerrogativa tecnológica para atender, por exemplo, ao viés social ante os desafios de acesso, recuperação e uso das informações nos cenários socioinstitucionais.

Diante da Era dos Dados e à luz dos Objetos Digitais, gerados como parte da revolução tecnológica, em que cabem os Objetos Digitais Imagéticos como parte do movimento de Dataficação no contexto de *Big Data*, é factível inferir: a Engenharia do Conhecimento, enquanto ciência, tem se ocupado de estratégias e metodologias adequadas para subsidiar os processos tecnológicos baseados em ferramentas, sobretudo em Inteligência Artificial, que reverberam no contexto social. Como já observado, isso não é uma realidade nos cenários da Ciência da Informação.

Portanto, é a Engenharia do Conhecimento, de modo multidisciplinar, quem preconiza, em suas metodologias, a Representação do Conhecimento. Mais especificamente, através das Ontologias, por elas figurarem como uma parte essencial à forma como o conhecimento é representado e estruturado.

As Ontologias como parte de Sistemas de Representação do Conhecimento tornam possível a organização das informações de maneira significativa, no sentido de que os sistemas possam entender e utilizar o conhecimento de maneira eficaz. É por meio da lógica, produto da modelagem do conhecimento, que regras de inferência são declaradas e se tornam inteligíveis a humanos e máquinas, para que sejam determinadas tarefas ou funções de um algoritmo, sendo as regras e raciocínio lógico abordagens essenciais para a tomada de decisões de maneira automatizada, ou seja, determinar como os Objetos Digitais devem se portar e agir de modo inteligente (Silva, 2023).

A transformação nos modelos de produção de registros imagéticos, que antes tinha como base a tecnologia analógica (preto e branco), passou a estar baseada no 0 e 1 (dígitos

binários). A realidade contemporânea, com o uso da tecnologia digital e dos computadores, eleva a complexidade dos registros, por sua dinâmica interativa e versátil quanto à estabilidade para gestão e organização, permitindo que cumpra uma demanda já estabelecida socialmente (Innarelli, 2023).

Segundo Silva (2023), a abordagem aplicada aos Objetos Digitais Imagéticos, entendidos como imagens estruturadas por dígitos binários e presentes em sistemas computacionais, atuam como documentos digitais e emergem como objeto de estudo para a Ciência da Informação, no âmbito de Sistemas de Representação do Conhecimento, adotando especificamente imagens em ambiente digital à luz da realidade social vigente.

Sob a égide da Engenharia do Conhecimento, a mineração, organização e acesso aos Objetos Digitais, mesmo que de forma disruptiva (para Ciência da Informação) já é realizada por mídias digitais e ferramentas (*tools*) de Inteligência Artificial, os quais criam, manipulam, convertem e habilitam padrões de imagens em textos. Alguns, inclusive, promovem a entrada e saída de mecanismos de busca cujos resultados obtidos são realizados por imagens, conforme exemplifica o caso do *Google Lens*. Acrescentamos que tal funcionalidade já é uma realidade em outras mídias digitais, como o *Pinterest*, havendo a adoção no *E-Commerce* internacional, nos seus principais setores: imóveis, arte, vestuário e utensílios domésticos.

A infraestrutura tecnológica que subsidia o desenvolvimento de ferramentas computacionais pode ser caracterizada por um repertório de processos que cabem às metodologias em Engenharia do Conhecimento, principalmente sob a perspectiva do campo da Inteligência Artificial, em que os dados existentes em um ambiente computacional passam a ser reconhecidos como agentes e trabalhados de modo sistematizado no desenvolvimento de sistemas inteligentes, especialmente aqueles que lidam com a manipulação e aplicação eficientes do conhecimento em diferentes domínios (Witten, *et al.* 2011).

O repertório de processos em Engenharia do Conhecimento pode ser caracterizado, por exemplo, nos estudos de Schreiber, *et al.* (2000) que é base seminal no campo, nos quais as etapas típicas para criar sistemas inteligentes capazes de manipular conhecimento baseado em Objetos Digitais incluem diversas atividades: análise do domínio (identificação de infraestrutura tecnológica existente, mineração de dados, mapeamento dos Objetos Digitais, análise estratégica do conhecimento em suas relações e restrições), captura do conhecimento (baseada em princípios da Gestão do Conhecimento para coleta e formalização do conhecimento a partir de metodologias como, por exemplo, entrevista com especialistas, grupo focal, mentoria etc.), modelagem do conhecimento (baseada em Organização do Conhecimento para processamento da linguagem natural e mediação aos agentes inteligentes),

desenvolvimento de sistemas (sob a perspectiva do campo da Inteligência Artificial no eixo de aprendizagem de máquina), implementação e monitoramento, assim como capacitação e suporte aos usuários.

Como prerrogativa de atuação no campo multidisciplinar da Engenharia do Conhecimento a metodologia *CommonKADS*, desenvolvida pelo consórcio europeu *Knowledge Acquisition and Design Structuring* (KADS), opera mediante um conglomerado de ferramentas, visando a captura, gestão e modelagem de conhecimento baseado em sistemas de bancos de dados digitais computadorizados (Van Harmelen; Fensel, 1995; Schreiber, *et al.* 2000). A metodologia *CommonKADS*, segundo Schreiber *et al.* (2000), é popularizada como principal recurso e específica modelos de requisitos ligados à Engenharia do Conhecimento, sendo três para delimitar a dimensão de contexto (Modelo da Organização, Modelo da Tarefa e Modelo do Agente), dois para dimensão de conceituação (Modelo do Conhecimento e Modelo de Comunicação) e um para dimensão do Artefato ou, como tratado aqui, Objeto Digital (Modelo do Projeto).

[...] Os modelos de conhecimento e comunicação produzem a descrição conceitual das funções e dados de resolução de problemas que devem ser tratados e entregues por um sistema de conhecimento. O design do modelo converte isso em uma especificação técnica que é a base para a implementação do sistema de software. [...] Assim, um projeto de conhecimento CommonKADS produz três tipos de produtos ou entregas: 1. Documentos modelo CommonKADS; 2. informações de gerenciamento de projetos; 3. *software* de sistema de conhecimento (Schreiber, *et al.* 2000, p. 18-19, tradução nossa).

É observado que a Engenharia do Conhecimento apreende uma série de processos que se consolidam científica e socialmente, além de compor oportunamente os principais nichos de atuação em nível mundial baseado em tecnologia com potencial para inovação tecnológica, o que inclui as plataformas e mídias digitais.

No quesito gestão, organização e desenvolvimento de infraestrutura baseada em Objetos Digitais (como especificidade pátria os Objetos Digitais Imagéticos, compreendidos como documentos digitais em natureza de dados algorítmicos), a Engenharia do Conhecimento tende a propor soluções, visando ferramentas para armazenamento, recuperação, acesso e uso mediante princípios como: as Ontologias em Organização do Conhecimento a partir do trato de imagens como dados que podem ser modelados, relacionados ou restritos sem necessariamente o aporte de teorias que fundamentam tais princípios na Ciência da Informação, principalmente os derivados do campo da Representação

da Informação e do Conhecimento com raiz nas Ciências Documentárias, matriz de processos clássicos vistos em Shera (1980).

Como observado em Silva (2023), a vantagem da Engenharia do Conhecimento sobre a Ciência da Informação pode ser situada no fato de que na primeira os agentes inteligentes são programados com lógica formal de ordem filosófica, sem que haja a necessidade de um compromisso ontológico no viés linguístico ou regras de representação da informação nos domínios canônicos, sendo o foco apresentar soluções práticas viáveis para funcionamento de sistemas de conhecimento na Era de Dados.

A Ciência da Informação possui restrições frente à Engenharia do Conhecimento no que se refere ao desenvolvimento de tecnologias computacionais para tratamento e recuperação dos Objetos Digitais Imagéticos, pois ainda não há esquemas de êxito na Organização do Conhecimento a partir do uso exclusivo de imagens nos Sistemas de Representação do Conhecimento que seguem o dogma do uso de conceitos, particularmente a unidade do referente, ou seja, a forma escrita (palavra).

As especificidades das abordagens das áreas de Engenharia do Conhecimento e Ciência da Informação aplicadas aos Objetos Digitais Imagéticos atuam de modo multidisciplinar, no sentido de metodizar o encadeamento e viabilizar a recuperação e difusão de dados digitais, em que a primeira está centrada aos ativos digitais, mineração e curadoria de dados e ferramentas disruptivas para tarefas computacionais voltadas à organização e acesso. Já a Ciência da Informação promove contribuições sobre teorias e processos sociais em informação, como o processamento de linguagem natural e sua tradução aplicada a esquemas inteligíveis a máquinas e desenvolvimento de algoritmos, sob a práxis da captura e modelização de domínios do conhecimento.

A investigação acerca dos Objetos Digitais Imagéticos de modo multidisciplinar nas distintas áreas de Engenharia do Conhecimento e Ciência da Informação possui a organização e acesso como pontos de intersecção. São elas fundamentais para reverberar o diálogo científico.

No entanto, isso vai muito mais além. Contam, ainda, com contribuições da Ciência de Dados e Computação, por meio do desenvolvimento de sistemas e linguagem de programação. E também podemos citar as interações profícuas com o campo da semiótica, no auxílio dos estudos de imagem aliados a processos de comunicação, que são estratégicos aos sistemas de negócios.

## 4 Considerações finais

Os cenários tecnológicos baseados em Inteligência Artificial e uso de mídias e plataformas digitais é um reflexo da Dataficação, na qual a Engenharia do Conhecimento tem atuado de modo aplicado, desde o surgimento dos primeiros computadores, com soluções algorítmicas para tornar sistemas baseados em dados e linguagem de programação cada vez mais avançados. Com isso, são supridas as necessidades de usuários do entretenimento ao comércio, em centros baseados em humanidades digitais, ao redor do mundo.

No contexto da Era de Dados, o comércio eletrônico e a cultura de mídias digitais impulsionam a geração de diferentes Objetos Digitais. De códigos fonte a documentos digitais como imagens, algo que configura um processo oportuno e necessário à Engenharia do Conhecimento, pois modela dados e implementa agentes inteligentes, os quais operam de modo automatizado mediante preceitos basilares da aprendizagem de máquina.

Como já analisado, os Objetos Digitais Imagéticos figuram como documentos baseados em dados e articulados algoritmicamente, em sistemas computacionais, para realização de tarefas pré-determinadas, como parte do movimento de Dataficação no contexto de *Big Data*. Eles emergem como uma oportunidade para áreas como a Engenharia do Conhecimento e possuem um repertório de processos orientados por aspectos multidisciplinares, dentre os quais foram destacados, também, àqueles que dialogam de maneira fronteira com saberes e fazeres da Ciência da Informação, a exemplo dos modelos semânticos e, principalmente, as Ontologias, no eixo de Gestão e Organização do Conhecimento.

O avanço das mídias digitais atua de modo propulsor à proliferação e consolidação da comunicação visual, em que se enquadram diferentes Objetos Digitais Imagéticos; imagens estáticas e em movimento de perspectiva nato-digitais ou representantes digitais (produto da digitalização).

Corroborando, assim, com o fenômeno de *Big Data*, cenário no qual a Organização do Conhecimento é pensado enquanto processo na Engenharia do Conhecimento subsidiado por preceitos comuns à Ciência da Informação. Opera, então, para metodizar o encadeamento e viabilizar a recuperação e difusão de dados digitais.

No desenvolvimento desta pesquisa, algumas das dificuldades encontradas foram: a escassez de fontes especializadas sobre Engenharia do Conhecimento em língua portuguesa, abordagens teóricas que estabeleçam o diálogo entre a Ciência da Informação e Engenharia do Conhecimento, ou mesmo acerca do contexto dos Objetos Digitais Imagéticos. Este aspecto



implica em limitações à discussão aqui apresentada, por não contar com múltiplos olhares científicos acerca da temática em questão.

Neste sentido, espera-se contribuir com o debate científico e despertar, de forma crítica, para a necessária discussão entre as áreas de Engenharia do Conhecimento e Ciência da Informação. Também é pretensão das autoras realizar novos estudos, que viabilizem o aprofundamento teórico da perspectiva ensaística aqui apresentada, bem como os alargamentos teórico e conceitual acerca da multidisciplinariedade da organização, recuperação e preservação dos Objetos Digitais Imagéticos.

Por fim, acredita-se na necessidade permanente dos diálogos entre teorias e práticas das áreas de Engenharia do Conhecimento e Ciência da Informação. Eles são ferramentas indispensáveis às reflexões nos cenários da tecnologia da informação, reverberando em contribuições potenciais à sociedade atual da Era de Dados. A partir do processo de Dataficação, é que cabem os aprofundamentos epistemológicos e pontos críticos, entre as similaridades, completudes e divergências teóricas e empíricas.

## Referências

- ARAÚJO, C. A. A. **O que é ciência da informação**. Belo Horizonte: KMA, 2018.
- BARDIN, L. **Análise de conteúdo**. São Paulo: Edições 70. 2011.
- BORKO, H. Information Science: What is it? **American Documentation**, v.19, n.1, p.3-5, jan. 1968. Disponível em: [https://edisciplinas.usp.br/pluginfile.php/1992827/mod\\_resource/content/1/Borko.pdf](https://edisciplinas.usp.br/pluginfile.php/1992827/mod_resource/content/1/Borko.pdf) Acesso em: 10 jan. 2024.
- CAPURRO, R. Epistemologia e ciência da informação. In: Encontro Nacional de Pesquisa em Ciência da Informação, 5., 2003, Belo Horizonte. **Anais**. Belo Horizonte: Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação e Biblioteconomia, 2003. Disponível em: <http://enancib.ppgci.ufba.br/artigos/GT1--231.pdf> Acesso: 10 jan. 2024.
- GIL, A. C. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.
- INNARELLI, H. C. **Cryptex da preservação digital**. 1 ed. Curitiba: Appris, 2023.
- MARQUES, A. J. **Metodología em ciencia dedutiva**. Juiz de Fora - MG. Ed. Gráfica Rio Branco. 1999.
- MINAYO, M. C. S. SANCHES, O. Métodos Qualitativos e Quantitativos: oposição ou complementaridade? **Caderno de Saúde Pública**, Rio de Janeiro, v. 9, n. 3, p. 239-262, jul./set. 1993. Disponível em: <https://www.scielo.br/j/csp/a/Bgpmz7T7cNv8K9Hg4J9fJDb/?format=pdf&lang=pt> Acesso em: 12 jan. 2024.

LÉVY, P. **Cibercultura**. São Paulo: Editora 34. 1999.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**. Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/22308> Acesso: 10 jan. 2024.

SCHREIBER, A. Th. *et. al.* **Knowledge engineering and management: the commonKADS methodology**. Massachusetts Institute of Technology. 2000.

SHERA, J. Sobre biblioteconomia, documentação e ciência da informação. In: GOMES, Hagar Espanha (org.) **Ciência da Informação ou informática?** Rio de Janeiro: Calunga, 1980.

SILVA, I. B. S. Recomendações para uma ontologia de domínio no âmbito de objetos digitais imagéticos à luz da Ciência da Informação. 2023. 112f. **Tese** (Doutorado em Ciência da Informação). Programa de Pós-Graduação em Ciência da Informação da Universidade Federal da Bahia (PPGCI/UFBA). Salvador, 2023. Disponível em: <https://repositorio.ufba.br/handle/ri/37408> Acesso em: 12 jan. 2024.

VAN HARMELEN, F. A. H., FENSEL, D. Formal Methods in Knowledge Engineering. **Knowledge Engineering Review**, 10(4), 345-360. 1995. Disponível em: <https://doi.org/10.1017/S0269888900007554> Acesso em: 13 jan. 2024.

WITTEN, I. H. EIBE, F. HALL, M. A. **Data mining: practical machine learning tools and techniques**. Ed. Morgan Kaufmann, 3rd ed. 2011.

## NOTAS

### CONTRIBUIÇÃO DE AUTORIA

**Concepção e elaboração do manuscrito:** I. B. S. Silva, H. F. Novo

**Coleta de dados:** I. B. S. Silva

**Análise de dados:** I. B. S. Silva

**Discussão dos resultados:** I. B. S. Silva, H. F. Novo

**Revisão e aprovação:** H. F. Novo

### DECLARAÇÃO DE DISPONIBILIDADE DOS DADOS

Todo o conjunto de dados que dá suporte aos resultados deste estudo foi publicado no próprio artigo.

### FINANCIAMENTO

Não se aplica

### CONFLITO DE INTERESSES

Não se aplica

### REVISÃO E NORMALIZAÇÃO

Os autores

### EDITOR RESPONSÁVEL

Patrícia Nascimento Silva (<https://orcid.org/0000-0002-2405-8536>)

### EQUIPE DE APOIO

Josiane Santos Lima (<https://orcid.org/0009-0001-2672-0351>)

### HISTÓRICO

Recebido em: 29-02-2024 – Aprovado em: 17-04-2024 – Publicado em: 03-05-2024.