

UNIVERSIDADE FEDERAL DE SÃO CARLOS (UFSCar)  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS (CECH)  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (PPGCI)

DENER CESAR FERREIRA LOPES

**GRAFOS DE CONHECIMENTO: PERSPECTIVAS E DESAFIOS PARA A  
ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO**

São Carlos  
2020

UNIVERSIDADE FEDERAL DE SÃO CARLOS (UFSCar)  
CENTRO DE EDUCAÇÃO E CIÊNCIAS HUMANAS (CECH)  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO (PPGCI)

DENER CESAR FERREIRA LOPES

**GRAFOS DE CONHECIMENTO: PERSPECTIVAS E DESAFIOS PARA A  
ORGANIZAÇÃO E REPRESENTAÇÃO DO CONHECIMENTO**

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Informação, como parte das exigências para obtenção do título de Mestre em Ciência da Informação.

Orientador: Prof. Dr. Rogério Ap. Sá Ramalho

São Carlos  
2020

**Lopes, Dener Cesar Ferreira**

Grafos de Conhecimento: perspectivas e desafios para a organização e representação do conhecimento / Dener Cesar Ferreira Lopes. – 2020. 78 f. : 30 cm.

Dissertação (mestrado)-Universidade Federal de São Carlos, campus São Carlos, São Carlos

Orientador: Rogério Aparecido Sá Ramalho

Banca examinadora: Rogério Aparecido Sá Ramalho, Roniberto Morato do Amaral, Vânia Mara Alves Lima, Bibliografia

1. Grafo de Conhecimento. 2. Organização do Conhecimento. Vocabulários controlados. 3. Tecnologias Semânticas. I. Orientador. II. Universidade Federal de São Carlos. III. Título.

Ficha catalográfica elaborada pelo Programa de Geração Automática da Secretaria Geral de Informática (SIn).

DADOS FORNECIDOS PELO(A) AUTOR(A)

Bibliotecário(a) Responsável: Ronildo Santos Prado – CRB/8 7325



UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Educação e Ciências Humanas  
Programa de Pós-Graduação em Ciência da Informação

Folha de Aprovação

Assinaturas dos membros da comissão examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato Dener Cesar Ferreira Lopes, realizada em 06/03/2020:

Prof. Dr. Rogério Aparecido Sá Ramalho  
UFSCar

Prof. Dr. Roniberto Morato do Amaral  
UFSCar

Profa. Dra. Vânia Mara Alves Lima  
USP

Certifico que a defesa realizou-se com a participação à distância do(s) membro(s) Vânia Mara Alves Lima e, depois das arguições e deliberações realizadas, o(s) participante(s) à distância está(ão) de acordo com o conteúdo do parecer da banca examinadora redigido neste relatório de defesa.

Prof. Dr. Rogério Aparecido Sá Ramalho

## DEDICATÓRIA

*Dedico à todos os meus antepassados pelas boas vibrações e energias libertadoras ora emanadas e por sua resiliente vontade e hercúlea predestinação de manterem-se sãos em terras e situações tão inóspitas possibilitando assim sua longeviva continuidade.*

## **AGRADECIMENTOS**

À *Natureza Divina*, por tudo que tem me proporcionado em termos de evolução, aperfeiçoamento, ciência e consciência. Por me permitir o “experienciar” da vida e seus mistérios através de seus constantes chamamentos.

Agradeço a minha família, que assimilou minha postura de mudança e compreendeu minha necessidade de resgate.

À minha querida esposa Andréa que soube abraçar a causa incentivando meus esforços e amparando minhas quedas, me apoiando a todo o momento. Sem esse amparo e compreensão, tudo seria muito mais difícil.

Ao meu pai, por sempre afirmar que a solução dos problemas está nos livros e, à minha mãe por sempre acreditar em mim.

Ao orientador e grande amigo Rogério Ramalho, que por vezes me abriu chances fundamentais para o ingresso na vida acadêmica a qual hoje se materializa após um longo período de sonho.

À Professora Doutora e amiga Zaira Zafalon, pela sua disponibilidade enquanto coordenadora do PPGCI, em me situar enquanto profissional de mercado no contexto acadêmico.

Aos professores do PPGCI UFSCAR que me deram todos os mapas e bússolas assertivos.

À UFSCAR e sua estrutura como um todo, pela calorosa e receptiva acolhida.

## RESUMO

LOPES, Dener Cesar Ferreira. **Grafos de Conhecimento**: perspectivas e desafios para a organização e representação do conhecimento: subtítulo. 2020. 78 f. . Dissertação (Mestrado em Ciência da Informação) – Universidade Federal de São Carlos, São Carlos, 2020.

**Palavras-chave:** Grafo de Conhecimento. Representação do Conhecimento. Organização da Informação. Tecnologias Semânticas.

Grafos de Conhecimento têm ganhado um crescente destaque na área acadêmica e corporativa, geralmente em estreita associação com tecnologias da Web, dados vinculados, big data e computação em nuvem. A presente pesquisa é de natureza qualitativa, tem caráter descritivo, e tem por objetivo realizar um estudo teórico sobre a utilização de Grafos de Conhecimento em ambientes organizacionais, identificando os principais desafios contemporâneos para a representação do conhecimento, e possibilidades de contribuição da área de Ciência da Informação na concepção e utilização deste instrumento. Como resultado foi identificado que os Grafos de Conhecimento são constituídos a partir de uma combinação de estruturas de representação de conhecimento, processos de gerenciamento de informação e algoritmos de busca, proporcionando uma visão didática decorrente da combinação de dados vinculados, processos de negócios eficientes e criativos, determinando uma arquitetura de soluções otimizada. Neste sentido, destaca-se a relevância da pesquisa considerando a pluralidade de definições do conceito de Grafo de Conhecimento e seu valor enquanto ferramenta, em função das necessidades corporativas, caracterizando-se como uma campo fértil para pesquisas na área de Ciência da Informação, a partir da interdisciplinaridade característica, favorecendo a padronização de definições e o desenvolvimento de instrumentos de representação fundamentadas no preceito de inovação e apoio para tomadas de decisão.

**Palavras-chave:** Grafos de Conhecimento. Representação do Conhecimento. Organização da Informação. Tecnologias Semânticas.

## ABSTRACT

LOPES, Dener Cesar Ferreira. **Knowledge Graphics**: perspectives and challenges for the organization and representation of knowledge. 2020. 78 f. Dissertation (Master in Information Science) - Federal University of São Carlos, São Carlos, 2020.

**Summary:** Knowledge Graphs have gained increasing prominence in the academic and corporate fields, generally in close association with Web technologies, linked data, big data and cloud computing. The present research is of a qualitative nature, has a descriptive character, and aims to carry out a theoretical study on the use of Knowledge Graphs in organizational environments, identifying the main contemporary challenges for the representation of knowledge, and possibilities of contribution from the Science area Information in the design and use of this instrument. As a result, it was identified that Knowledge Graphs are constituted from a combination of knowledge representation structures, information management processes and search algorithms, providing a didactic vision resulting from the combination of linked data, efficient and creative business processes, determining an optimized solution architecture. In this sense, the relevance of the research is highlighted considering the plurality of definitions of the concept of Knowledge Graph and its value as a tool, according to the corporate needs, being characterized as a fertile field for research in the area of Information Science, the from the characteristic interdisciplinarity, favoring the standardization of definitions and the development of representation instruments based on the precept of innovation and support for decision making.

**Keywords:** Knowledge Graph. Knowledge Representation. Information Organization. Semantic Technologies.

## LISTA DE FIGURAS

Figura 1 - Crescimento das literaturas científicas	19
Figura 2 - Diferenças quando ao grau de complexidade de cada tipo de vocabulário	42
Figura 3 - Processo de descoberta de conhecimento em bases de dados	52
Figura 4 - Estrutura de um Grafo de Conhecimento	54
Figura 5- Arquitetura para Grafo de Conhecimento em organizações	59
Figura 6 - Exemplos de perguntas respondidas pelo Google Search Engine	64
Figura 6.1 - Exemplos de perguntas respondidas pelo Google Search Engine	64
Figura 6.2 - Exemplos de perguntas respondidas pelo Google Search Engine	65

## LISTA DE TABELAS

Quadro 1 – Definições de Grafos de Conhecimento	51
Tabela 2 - Comparação entre Grafos de Conhecimento	53

## **LISTA DE ABREVIATURAS E SIGLAS**

ABNT Associação Brasileira de Normas Técnicas

ANSI American National Standard Institute

API Application Programming Interface

CDD Classificação Decimal de Dewey

CDU Classificação Decimal Universal

HTML HyperText Markup Language

IBICT Instituto Brasileiro de Informação em Ciência e Tecnologia

IBM International Business Machines Corporation

IDC International Data Corporation

IOT Internet of Things

IRI Identificador de Recursos Internacionalizados

ISO International Organization for Standardization

KO Knowledge Organization

KOS Knowledge Organization System

MIT Massachusetts Institute of Technology

NISO National Information Standards Organization

OC ORGANIZAÇÃO DO CONHECIMENTO

OWL Web Ontology Language

PME Pequenas e Médias empresas

RDF Resource Description Framework

SGML Standard Generalized Markup Language

SKOS Simple Knowledge Organization System

SOC Sistema de ORGANIZAÇÃO DO CONHECIMENTO

SRI Sistema de Recuperação de Informação

TCT Teoria Comunicativa da Terminologia

TGT Teoria Geral da Terminologia

TIC Tecnologias de Informação e Comunicação.

XML eXtensible Markup Language

W3C World Wide Web Consortium

WWW World Wide Web

UFSCar Universidade Federal de São Carlos

URI Uniform Resource Identifier

## Sumário

<b>1.</b>	<b>INTRODUÇÃO .....</b>	<b>15</b>
<b>1.1.</b>	<b>PROBLEMA DE PESQUISA .....</b>	<b>21</b>
<b>1.2.</b>	<b>OBJETIVOS .....</b>	<b>22</b>
<b>1.3.</b>	<b>JUSTIFICATIVA .....</b>	<b>23</b>
<b>1.4.</b>	<b>PROCEDIMENTOS METODOLÓGICOS .....</b>	<b>24</b>
<b>1.5.</b>	<b>ESTRUTURA DA PESQUISA .....</b>	<b>24</b>
<b>2.</b>	<b>CIÊNCIA DA INFORMAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO .....</b>	<b>27</b>
<b>2.1.</b>	<b>CONTEXTUALIZANDO A CIÊNCIA DA INFORMAÇÃO.....</b>	<b>28</b>
<b>2.2.</b>	<b>A ÁREA DE ORGANIZAÇÃO DO CONHECIMENTO .....</b>	<b>32</b>
<b>3.</b>	<b>SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO .....</b>	<b>39</b>
<b>3.1.</b>	<b>VOCABULÁRIOS CONTROLADOS .....</b>	<b>40</b>
<b>3.2.</b>	<b>ONTOLOGIAS E REDES SEMÂNTICAS.....</b>	<b>43</b>
<b>3.3.</b>	<b>GRAFOS DE CONHECIMENTO .....</b>	<b>45</b>
<b>4.</b>	<b>PERSPECTIVAS DE UTILIZAÇÃO DE GRAFOS DO CONHECIMENTO EM AMBIENTES CORPORATIVOS .....</b>	<b>58</b>
<b>5.</b>	<b>CONSIDERAÇÕES FINAIS .....</b>	<b>66</b>
<b>6.</b>	<b>REFERENCIAS.....</b>	<b>71</b>

## 1. INTRODUÇÃO

Após a publicação dos livros *Choque do Futuro* (1970) e *A Terceira Onda* (1980) Alvin Toffler, escritor norte-americano, participou de uma série de entrevistas sugeridas por membros da South End Press<sup>1</sup>, onde as transcrições originaram o livro *Previsões e Premissas*, de 1983, uma das questões nos salta aos olhos. Segundo Toffler (1983, p. 124),

Se a Segunda Onda nos deu um “proletariado”, poderíamos dizer que a Terceira está produzindo um “cognitariado”, um grupo que se assenta nos conhecimentos, no emprego da mente, não do músculo. O cognitariado possui informações organizadas, imaginação e outras qualidades culturais essenciais para a produção. Possui os meios de produção de mais informações. Possui o que pode ser considerado ou como matéria prima essencial ou, alternativamente, um jogo de ferramentas mentais.

A decorrente previsibilidade citada anteriormente denota o acelerado desenvolvimento da tecnologia da informação e do conhecimento, influenciando áreas de pesquisa, atividades, reorganizando profissões e profissionais a partir da imposição de um novo ritmo que impacta os fluxos econômicos, produtivos, sociais, corporativos e financeiros da sociedade.

As volumosas massas informacionais são exponencialmente geradas e trafegam em velocidades diretamente proporcionais aos seus propósitos e nos atingem intimamente, determinando vieses e vivências que nos tornam portadores diretamente responsáveis, ou não, pelo seu ciclo evolutivo.

De acordo com Alvares (2012, 2012, p. 23),

conhecimento e informação são termos de difícil conceituação devido à amplitude semântica e às diversas perspectivas de análise, domínios e concepções de cada área, como já foi problematizado por alguns autores, dentre eles, Le Coadic (2004). Na ciência da informação, tais reflexos são frutos de reflexões teóricas e por vezes controversas.

Ainda de acordo com Alvares (2011), a informação obtida por um indivíduo para se transformar em conhecimento, dialoga com sua cultura, seus valores e princípios, do seu modo de ser e sua maneira de compreender o mundo. O conhecimento, nesse caso, é subjetivo (inerente ao sujeito), mas ao mesmo tempo social, pois o ser humano interage com o mundo que o circunda, modificando-o e sendo por ele modificado. Nem toda informação existente

---

<sup>1</sup> A South End Press é uma editora de livros independente, sem fins lucrativos e administrada coletivamente, com mais de 250 títulos impressos. Desde a sua fundação em 1977, eles atendem às necessidades de leitores que estão explorando ou já estão comprometidos com a política de mudanças sociais radicais. Seu objetivo é publicar livros que incentivem o pensamento crítico e a ação construtiva sobre os principais problemas políticos, culturais, sociais, econômicos e ecológicos que moldam a vida nos Estados Unidos e no mundo. Eles esperam oferecer um fórum para uma ampla variedade de movimentos sociais democráticos e oferecer uma alternativa aos produtos da publicação corporativa. Disponível em < <http://southendpress.org/>>. Acesso em 25 jan. 2020.

em um documento vai se transformar em conhecimento, pois que aprende precisa ter os elementos fundamentais para a decodificação da informação, ou seja, fazer a correlação dessa informação com as estruturas mentais e conhecimentos correlatos mínimos que possibilitarão o entendimento e, se for o caso, a geração de novos conhecimentos.

Conforme afirma Furnival (2005, p.16, 2005),

na profusa literatura sobre questões contemporâneas é facilmente perceptível a afirmação de que vivemos, desde a segunda metade do século XX, na chamada “sociedade da informação” ou “sociedade do conhecimento”. Como uma decorrência dessa característica, pode ser mencionada a intensificação da circulação de bens intangíveis, como a informação e o conhecimento, assim como o aumento da capacidade de geração e uso desse bem como insumo-chave para o desenvolvimento das atividades econômicas e de ações sociais e políticas.

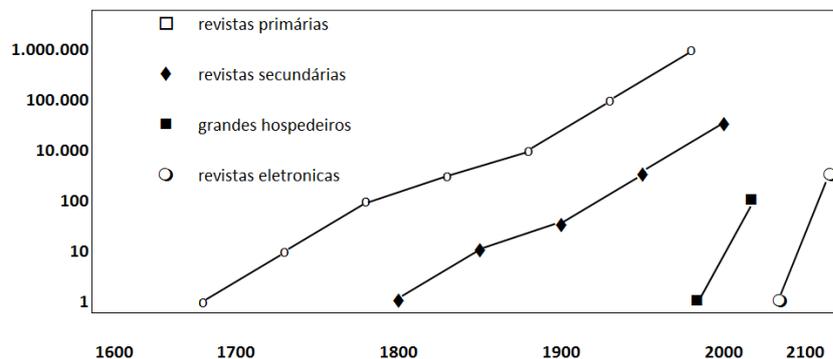
A sociedade da informação traz um novo cenário tecnológico e informacional, objetivando e fomentando uma maior consciência política dos cidadãos e provocando o setor produtivo sobre suas responsabilidades inerentes ao contexto evolutivo de suas ações e consequentes impactos sociais, ambientais, entre outros.

Sobre a explosão quantitativa da informação, de acordo com Le Coadic (2004, p.5),

o progresso técnico e social no poder criativo da linguagem e do raciocínio lógico que daí resulta, pode-se compreender a importância da comunicação verbal da informação. Com o advento da escrita a comunicação passou de oral para escrita. Isso teve como consequência, por um baixo custo energético, multiplicar a informação (cópia de manuscritos, imprensa, fotocópia) e armazená-la, permitindo assim exteriorizar, primeiramente nas bibliotecas, uma das funções do cérebro humano, que é a memória. Essas operações de multiplicação e armazenamento explica boa parte do que se costumou chamar de explosão da informação (mais exatamente explosão da quantidade de informações): um crescimento que obedece a uma lei de tipo exponencial.

Nesta perspectiva, pode-se citar como exemplo de progresso técnico e social o crescimento exponencial do número de revistas científicas primárias, conforme pode ser observado na figura 1.

**Figura 1.** Crescimento das literaturas científicas.



Adaptado de Le Coadic (2004)

Segundo Le Coadic (2004) as revistas de resumo (chamadas revistas secundárias), criadas para facilitar o acesso às revistas primárias, tiveram igual evolução. Sua informatização, que está na origem dos bancos de “dados”, apenas desloca o problema, pois também passaram pelo mesmo tipo de crescimento, o que implica o surgimento de hospedeiros cada vez mais importantes, verdadeiros supermercados da informação.

Ao longo dos últimos anos o crescimento exponencial da quantidade de informações disponíveis, decorrente de fatores como: o desenvolvimento de novas tecnologias, usos e costumes, especialização das redes estáticas e móveis, popularização de repositórios remotos e das infraestruturas pré-definidas que abarcam contextos corporativos, têm acarretado um aumento significativo da complexidade dos processos de organização do conhecimento em ambientes organizacionais.

Tomando-se por escopo de atuação da ciência da informação, os Sistemas de Organização do Conhecimento ou SOC, são representações do conhecimento, que por sua vez são modelos de abstração do mundo real, construídos com determinada finalidade. São sistemas de classificação, dicionários, ontologias, taxonomias, tesouros, redes semânticas, entre outros, de acordo com Alvares (2012).

Os SOC são representações do conhecimento, que por sua vez são abstrações do mundo real, construídos para determinada finalidade, conforme Alvares (2012).

HØjarland (2008) define sistemas de organização do conhecimento como ferramentas que apresentam a interpretação organizada de estruturas do conhecimento, também chamadas ferramentas semânticas. Contemplam, essencialmente, conceitos e suas relações semânticas.

Alvares (2012) aponta relevante função dos SOC em bibliotecas digitais: a de estabelecer elos entre recursos digitais e entre esses e objetos físicos. Os registros podem ser localizados diretamente por meio da navegação nos SOC ou indiretamente, por meio de uma chave de busca que pode ser usada para acessar outra fonte de informação.

Com a popularização das denominadas máquinas inteligentes<sup>2</sup> desenvolvidas a partir de treinamentos condicionados, pelas transações globais que determinam os vários movimentos de mercado e as características de prontidão e assertividade dos contextos *on* e *off line*, percebe-se claramente a emergente necessidade de fluência do contexto informacional.

Decorrente da quantidade crescente de dados, a agregação de conhecimento, a representação e o raciocínio aplicado são altamente fundamentais para as empresas, de modo que a representação concisa do conhecimento se torna fundamental para favorecer a tomada de decisões.

Neste contexto, segundo PEREZ et al. (2017) o Grafo de Conhecimento tem sido cada vez mais utilizado em ambientes organizacionais e de pesquisa para determinar instrumentos de representação do conhecimento conjugados às tendências de pesquisa envolvendo: Dados Vinculados, Web Semântica, Análise de Dados, Computação em Nuvem e Ciência de Dados.

Considerando, em especial, as decorrências após recentes crises financeiras, as empresas procuram processos de negócios muito mais eficientes e criativos, a fim de colocar melhores soluções no mercado em menos tempo e menos custo. Há uma impressão geral de que a comunicação e a colaboração, especialmente conjugadas às abordagens da Web 2.0<sup>3</sup> nas empresas e nos ecossistemas, podem impulsionar o processo de inovação com impactos positivos nos indicadores de negócios.

---

<sup>2</sup> Conforme Frazão e Mulholland (2019), apesar de a Inteligência Artificial (IA) ser um assunto pesquisado há décadas, ela continua sendo um dos tópicos mais explorados na Ciência da Computação. Isso se deve ao fato de que a IA aborda um amplo aspecto de investigações, que vão desde questões conceituais, como “Máquinas podem pensar?”, a soluções práticas tais como algoritmos de busca aplicados a jogos.

<sup>3</sup> Tim O'Reilly, o precursor do uso do termo em seu artigo de conceitualização (e também de defesa) do termo Web 2.0, define que: "Web 2.0 é a mudança para uma internet como plataforma, e um entendimento das regras para obter sucesso nesta nova plataforma. Entre outras, a regra mais importante é desenvolver aplicativos que aproveitem os efeitos de rede para se tornarem melhores quanto mais são usados pelas pessoas, aproveitando a inteligência coletiva" Disponível em < <https://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html> > Acesso em 25 Jan 2020.

Para Chesbrough (2012), a Inovação Aberta no contexto da Empresa 2.0 é um dos paradigmas mais populares para melhorar os processos de inovação das empresas, com base na criação e desenvolvimento colaborativos de ideias e produtos. A principal característica desse novo paradigma é que o conhecimento é explorado de forma colaborativa, fluindo não apenas entre fontes internas, por exemplo, Departamentos de Desenvolvimento de Produtos, mas também entre externos, como funcionários, clientes, parceiros etc. Nesse cenário, segundo PEREZ et al. (2017), podem-se destacar duas formas de utilizações principais de Grafos de Conhecimento em ambientes corporativos:

- Apoiar a contextualização semântica de tarefas relacionadas ao conteúdo envolvendo indivíduos e funções;
- Ajudar na descoberta de relações entre comunidades de funcionários, clientes e fornecedores, com conhecimentos e interesses compartilhados.

De acordo com Kelley (2007) a introdução do paradigma de Inovação Aberta<sup>4</sup> em uma empresa envolve não apenas uma modificação dos processos de inovação corporativa, mas também uma mudança cultural que requer suporte de uma infraestrutura tecnológica avançada.

O conhecimento corporativo, segundo PEREZ et al. (2017), deve ser explicitado, trocado e compartilhado entre os participantes; portanto, são necessárias ferramentas para gerenciamento do conhecimento, suporte à análise e estruturação da informação para tornar essas tarefas acessíveis e o conhecimento disponível para todos os atores envolvidos. Além disso, as ferramentas que suportam o processo de inovação precisam fornecer um alto grau de interatividade, conectividade e compartilhamento.

Em um cenário em que o trabalho colaborativo não é suportado e os membros de uma comunidade mal conseguem interagir entre si, as soluções para os problemas cotidianos e os problemas organizacionais contam com a iniciativa de um indivíduo específico para aquele trabalho. Para PEREZ et al. (2017), inovação e gestão de áreas como a de Desenvolvimento de Produtos são processos complexos para os quais colaboração e comunicação são fundamentais.

---

<sup>4</sup> Termo criado por Henry Chesbrough, professor e diretor executivo no Centro de Inovação Aberta da Universidade de Berkeley e chairman do Centro de Open Innovation - Brasil.

Inovação e gestão, conjuntamente, implicam na criação, reconhecimento e articulação de oportunidades, que precisam ser evoluídas para uma proposição de negócios em um estágio subsequente. Interatividade, conectividade e compartilhamento são os recursos a serem considerados ao projetar uma estrutura tecnológica para apoiar inovações colaborativas.

Todas essas características podem ser identificadas nos ambientes da Empresa 2.0<sup>5</sup>. No entanto, as ferramentas da Empresa 2.0, ainda segundo PEREZ et al. (2017), não fornecem modelos formais que são usados para criar sistemas complexos que gerenciam consideráveis quantidades de informações. Essa desvantagem, segundo PEREZ et al. (2017), pode ser superada com a incorporação dos Grafos de Conhecimento corporativo que introduzem representações interligadas e lidas por programas de computador. As plataformas de inovação aberta alavancam o conceito de um Grafo de Conhecimento corporativo para relacionar pessoas, interesses e ideias em um ambiente de gerenciamento de conhecimento corporativo em todos os setores, envolvendo funcionários, clientes e outras partes interessadas.

De acordo com PEREZ et al. (2017), o impacto dos Grafos de Conhecimento através de sua aplicação em inovação aberta é ilustrado por sua adoção em corporações pertencentes a vários setores, como bancos, energia e telecomunicações, com empresas como Bankinter, Repsol e Telefonica, que se posicionaram na vanguarda desses esforços. O que todos esses esforços têm em comum é a necessidade de conectar ideias e pessoas inovadoras, a fim de orquestrar um ecossistema de inovação saudável, enfrentando vários desafios, como:

- Manipular as informações criadas por milhares de funcionários;
- Avaliar suas ideias com eficiência;
- Redução de falsos positivos (ideias que chegam ao mercado e falham) e falsos negativos (ideias valiosas que são rejeitadas antes mesmo que cheguem ao mercado);
- Estimular a comunicação entre pessoas localizadas em todo o mundo, em diferentes idiomas.

De acordo com Paulheim (2016), os Grafos de Conhecimento na Web são uma espinha dorsal de muitos sistemas de informação que requerem acesso ao conhecimento estruturado, seja específico ou independente do domínio. A ideia de alimentar sistemas

---

<sup>5</sup> O termo “Empresa 2.0” foi dado pelo pesquisador do Center for Digital Business do MIT, Andrew McAfee, para descrever a forma como as empresas estão utilizando um novo conjunto de tecnologias (blogs, SRS, Twitter, Wikis, entre outras.) Disponível em <[https://seer.uscs.edu.br/index.php/revista\\_gestao/article/download/1205/1000](https://seer.uscs.edu.br/index.php/revista_gestao/article/download/1205/1000)> Acesso em 22 Jan 2020.

inteligentes agentes com conhecimento geral e formalizado. Assim sendo, remontamos à pesquisa clássica de Inteligência Artificial dos anos 80. Mais recentemente, com o advento de fontes de dados vinculados e de dados abertos como o DBpedia<sup>6</sup>, e pelo anúncio do Google pelo Google Grafo de Conhecimento em 2012, representações do conhecimento do mundo em geral com grafos despertam a atenção às oportunidades notadamente evolutivas quando avaliamos impactos de organização e representação a partir de silos de informações alcunhados Big Data.

A proposta desta pesquisa é traçar paralelos evolutivos relacionados ao surgimento do conceito de Grafo de Conhecimento e sua interação com a área de Ciência da Informação. Foi realizada uma pesquisa bibliográfica sobre os fundamentos que norteiam Grafos de Conhecimento e a arquitetura que os constitui, a relação com contextos organizacionais e cases de sucesso identificados na literatura. Em um segundo momento são apresentadas as formas de construção, uma síntese de sua arquitetura lógica e a estrutura de consumo da informação representada, bem como suas aplicações em contextos corporativos, histórias de sucesso e um olhar de expectativas para o futuro quanto aos benefícios e valor agregado tal instrumento representa na busca pela liderança competitiva. É nessa jornada que se fixa também, entre outros contextos, a importância do conhecimento e informação na era digital e, as decorrentes alianças impulsionadas pelo conhecimento e o desenvolvimento de sua representação baseada em colaboração e co-criação.

Ao longo da pesquisa foram analisados diferentes tipos de tecnologias e estruturas de representação, variando desde linguagens básicas de representação como RDF à estruturas mais avançadas como ontologias e redes neurais, favorecendo um melhor entendimento para a compreensão de Grafos de Conhecimento.

## **1.1. PROBLEMA DE PESQUISA**

Desde o anúncio feito pelo Google da criação de seu Grafo de Conhecimento, em 2012, a adesão à tecnologia por várias entidades tem repercutido em uma variedade de interpretações, por vezes condicionadas ou pela circunstância a qual é utilizado, ou pela área

---

<sup>6</sup> A DBpedia é uma base de conhecimento aberta, gratuita e abrangente, constantemente aprimorada e ampliada por uma grande comunidade global. Disponível em < <https://wiki.dbpedia.org/> > Acesso em 25 Jan 2020.

de conhecimento que o utiliza para suas finalidades, o tem dificultado consideravelmente a equalização de um entendimento comum dos Grafos de Conhecimento.

A realidade que se tornou um desafio em vários segmentos de mercado é a manipulação e a decorrente necessidade de representação do conhecimento envolvido em muitos dos setores empresariais de grande porte. A postura corporativa admite novos contornos onde tecnologias passam a ditar uma nova ordem na condução de negócios que já é base para uma evolução comercial de muitas empresas. Muitas das tendências tecnológicas como a transformação digital, o Big Data, o IOT, a Ciência de Dados, a Inteligência Artificial, e todas as estratégias e metodologias de vanguarda oferecem uma gama de alternativas as quais podem ser o divisor de águas entre o sucesso e o fracasso de uma corporação quando se trata de usabilidade versus resultados. A observação em questão é um alerta para algo bastante pertinente: como veremos a seguir, o Grafo de Conhecimento compõe-se de cada uma dessas tecnologias as quais são evidentes em sua arquitetura.

As respostas por vezes esperadas em uma solução ou direção a ser adotada, podem ser diversas do que se pretende caso a conceituação do que se tem como expectativas, conjugadas à aplicação tácita da solução, pode trazer resultados diversos do que se espera, caso não haja antecipadamente um alinhamento quanto ao conceito que se pretende a aplicar.

Assim, considerando a diversidade de necessidades corporativas e mercadológicas que permeiam as atividades econômicas, que permeiam a utilização de Grafos de Conhecimento e os instrumentos de Representação do Conhecimento identificados na área de Ciência da Informação pode-se questionar: Quais os reflexos da utilização de Grafos de Conhecimento na representação do conhecimento em ambientes organizacionais e quais as contribuições da área de Ciência da Informação neste novo cenário?

## **1.2. OBJETIVOS**

O presente trabalho tem como objetivo geral realizar um estudo teórico sobre a utilização de Grafos de Conhecimento em ambientes organizacionais, identificando os principais desafios contemporâneos para a representação do conhecimento, e verificar possibilidades de contribuição da subárea de organização da Ciência da Informação na concepção e utilização deste instrumento.

Para atingir o objetivo da pesquisa, delinearão-se os seguintes objetivos específicos:

- Realizar uma revisão de literatura sobre Sistemas de Organização do Conhecimento e Tecnologias de Representação;
- Analisar como os Grafos de Conhecimento podem ser utilizados, identificando suas principais características e relações com as novas tecnologias de representação.

### **1.3. JUSTIFICATIVA**

Ao nos aprofundamos na análise dos Grafos de Conhecimento, verificamos a quantidade de compartilhamento de informações que são possíveis neste contexto. É entendido também que a tendência de utilização que permeia os benefícios do compartilhamento de informações é extrapolar os limites das corporações, conectando ativos de dados de empresas ao longo de suas cadeias de valor. Uma das justificativas que entendemos ser pertinente na pesquisa é o estabelecimento de regramentos ou normativas que possam balizar acordos e utilização quanto às entidades que participem direta ou indiretamente de compartilhamentos informacionais.

Concomitante à previsibilidade da expansão da utilização dos Grafos de Conhecimento bem como para a criação de aplicativos mais úteis na agricultura, telecomunicações, finanças, saúde e outros setores, os benefícios oferecidos são muitos: integração e fusão de dados, melhores processos de curadoria de dados, manutenção de dados a custos reduzidos e melhores informações sobre as informações que estão sendo tratadas. O estabelecimento de parcerias de dados, criando um ecossistema rico com valor adicional, tanto para as empresas quanto para a sociedade como um todo.

A comunidade científica deve ter seu posicionamento melhor amparado em seu arcabouço teórico, envolvendo-se nas necessidades advindas na orientação de padrões de práticas e observâncias de componentes reutilizáveis em todos os níveis de desenho de soluções que envolvam linguagens, sintaxes, semânticas, padrões de conhecimento, padrões de consulta, padrões de raciocínio. Com isso, o objetivo de equalização de conceitos é sanado e ambiguidades são mitigadas tanto no processo de arquitetura quanto na aplicação dos Grafos de Conhecimento.

## **1.4. PROCEDIMENTOS METODOLÓGICOS**

Para Lakatos (2007), todo projeto de pesquisa deve conter as premissas ou pressupostos teóricos sobre os quais o pesquisador (o coordenador e os principais elementos de sua equipe) fundamentará sua interpretação. Trata-se de uma pesquisa teórica que se propõe a direcionar uma melhor compreensão do conceito e aplicação do Grafo de Conhecimento em ambientes organizacionais. Possui abordagem qualitativa, pois busca apresentar características voltadas à descrição, compreensão e explicação de relações de um determinado fenômeno de modo a considerar o caráter interativo entre os objetivos definidos e as orientações teóricas da pesquisa.

A pesquisa tem caráter descritivo, de coleta bibliográfica (revisão bibliográfica, pesquisa bibliográfica( - expressão de busca - fontes de informações - critérios de seleção - atividade de análise - análise de conteúdo - apresentação dos resultados) e de natureza qualitativa. Caracteriza-se, também, como pesquisa de natureza aplicada, uma vez que “objetiva gerar conhecimentos para aplicação prática, dirigidos à solução de problemas específicos” (SILVEIRA; CÓRDOVA, 2009, p.35).

O caráter exploratório da pesquisa, segundo Cervo e Bervian (2003), consiste na busca de informações sobre o assunto a ser estudado, considerando os diversos aspectos de um problema a fim de solucioná-lo. Busca tornar o problema mais explícito a partir de hipóteses, o que exige descrição de fatos observados (GIL, 2007; TRIVIÑOS, 1987). Segundo Köche (2002, p. 126) estudos exploratórios desencadeiam “[...] um processo de investigação que identifique a natureza do fenômeno e aponte as características essenciais das variáveis que se quer estudar”. Além disso, o caráter descritivo da pesquisa permitirá descrever as características do objeto de estudo investigado (COSTA, COSTA, 2001).

## **1.5. ESTRUTURA DA PESQUISA**

As seções que seguem, iniciam-se com a visão e posicionamento de conceitos que consubstanciam a Ciência da Informação, seu papel enquanto área interdisciplinar e sua importância enquanto ciência, conforme preceitua a seção 2. Ciência da Informação e Organização do Conhecimento. Nela conjugam-se conceitos voltados para a organização do conhecimento enquanto prática e disciplina. No que se refere à organização e representação

da informação, há, questões ainda sem resposta as quais a Ciência da Informação se incumbem de investigar-las.

Em sentido correlato, a seção 2.1 Contextualizando a Ciência da Informação nos remete à preceitos que identificam a trajetória de atuação da C.I. e reforça seu papel mediador junto às demais ciências tais como, a Matemática, Lógica, Linguística, Psicologia, Ciência da Computação, Engenharia da Produção, Artes Gráficas, Comunicação, Biblioteconomia, Administração, e outros campos científicos semelhantes através de seu aporte teórico e metodológico.

A seção 2.2 A área de Organização do Conhecimento traz consigo a necessidade de marcação de território de atuação concomitante às demais ciências e a relevância da pesquisa científica voltada para a organização e representação do conhecimento, dando insumos ao que contextualiza a seção 2.3 Organização e Representação do Conhecimento enquanto campos de pesquisa científica desenvolvendo regras e padrões, incluindo sistemas de classificação, listas de títulos de assuntos, tesouro e outras formas de metadados.

Na seção 3. Sistemas de Organização do Conhecimento são abordados os SOCs bem como algumas de suas funções, complexidades e objetivos. Seguindo adiante, temos a seção

3.1. Vocabulários Controlados a qual aborda as representações de maneira sistêmica e funcional e sua conjugação com recursos computacionais demonstrando visões que vão da linguagem às ferramentas.

Na seção 3.2. Ontologias e Redes Semânticas, além de conceitos temos a evidente interdisciplinaridade que orbita na ciência da informação compartilhando constructos tanto para o campo das ontologias quanto para o universo de redes.

Na seção 3.3. Grafos de Conhecimento há um delineamento histórico, conceitual e arquitetônico enfatizando sua importância enquanto ferramenta de representação do conhecimento bem como mediador de volumes informacionais em tempo real.

Na seção 4. Perspectivas de Utilização de Grafos do Conhecimento em Ambientes Corporativos, uma visão baseada em novas tecnologias bem como comportamentos reformulados através de mudanças nos fluxos informacionais, tornam os Grafos de Conhecimento ferramentas poderosas que determinam caminhos para corporações que, por

consequente, dão novos contornos ao formato representativo do conhecimento, recuperação, armazenamento e aprendizado.

Por fim, temos a seção 5. Considerações finais, a qual cria um contexto de remissão ao conteúdo abordado, e sugere visões futuras quanto aos próximos passos decorrentes de uma tecnologia que tende a criar ecossistemas voltados para uma economia informacional emergente.

## 2. CIÊNCIA DA INFORMAÇÃO E ORGANIZAÇÃO DO CONHECIMENTO

Em meados do século XX o aparecimento da Ciência da Informação está intimamente ligado às mudanças no campo da Biblioteconomia/ /Documentação, por força da tecnologia, bem como a percepção de que novos desafios se começavam a colocar aos profissionais da área sujeitos a um novo paradigma, dinâmico, informacional e pós-custodial. Embora o fato da construção da ciência da Informação ter tido um percurso associado à evolução da informação científico-técnica, cujas origens remontam às ideias defendidas por Paul Otlet e Henri la Fontaine, a verdade é que, assistimos, na generalidade dos serviços de informação, a um uso crescente das novas tecnologias da informação e da comunicação (TIC), o que causou, irreversivelmente, uma crise paradigmática e a necessidade de uma mudança.

De acordo com Ribeiro (2005), a simbiose das TIC com o fenômeno da informação, sobretudo nas últimas duas décadas em que entramos definitivamente na "era digital", passou a ser uma realidade nova em que a tecnologia já não é, como antes, apenas um recurso ou uma ferramenta que ajuda a operacionalizar o tratamento e a recuperação da informação, mas tornou-se indissociável desta, em todo o seu ciclo vital: produção, tratamento, uso e armazenamento.

Uma evidência é a Internet. As possibilidades de consultas, feitas em segundos sobre volumes de informação, mudou consideravelmente os comportamentos de pesquisa, os perfis dos usuários e as necessidades de informação, agora à escala muito maior.

No que se refere à organização e representação da informação, há, questões ainda sem uma total resposta: as facilidades tecnológicas introduziram dinamismo e rapidez nos procedimentos. Os modelos tradicionais foram ajustados por efeito do uso das tecnologias.

Para Santos (2015, p.2)

Inserida no domínio da Ciência da Informação, como disciplina estrutural da área de aplicação Biblioteconomia, se apropria da fundamentação teórica da grande área para a personalização de arquiteturas informacionais para o uso estratégico das tecnologias de informação e comunicação (TIC) para a representação de recursos no favorecimento do acesso às informações, ocasionando a evolução de sistemas, técnicas e máquinas para o tratamento da informação, como os estudos teóricos e experimentais sobre a construção de registros descritivos, a estrutura do conhecimento e seus registros; a modelagem de catálogos; os usuários da informação; a interação homem---computador; a economia, impacto e valor da informação; entre outros.

Para Ribeiro (2005), representar informação significa criar "imagens", não exatas e integrais (isso seria uma reprodução *stricto sensu*), mas suficientemente rigorosas para tornar possível uma identificação inequívoca dos objetos representados. Para esse efeito usa-se a chamada "descrição" ou "representação descritiva" que, hoje em dia, começa a ser designada por meta-informação (informação sobre a própria informação). Sendo esta a noção de "representação", é óbvio que, de um ponto de vista prático, a operação só pode ser bem sucedida se houver um conhecimento prévio daquilo que vai ser descrito.

O objetivo final de um produto de informação ou de um sistema de informação deve ser pensado em termos dos usos dados à informação e dos efeitos resultantes desses usos nas atividades dos usuários. A função mais importante do produto ou do sistema é, portanto, a forma com que a informação modifica a realização dessas atividades. Por causa disso, devem ser "orientadas para o usuário" (COADIC, 2004).

## 2.1. CONTEXTUALIZANDO A CIÊNCIA DA INFORMAÇÃO

Em sua definição, Borko (1968) afirma que,

Ciência da Informação é a disciplina que investiga as propriedades e o comportamento informacional, as forças que governam os fluxos de informação, e os significados do processamento da informação, visando à acessibilidade e a usabilidade ótima. A Ciência da Informação está preocupada com o corpo de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação, e utilização da informação. Isto inclui a pesquisa sobre a representação da informação em ambos os sistemas, tanto naturais quanto artificiais, o uso de códigos para a transmissão eficiente da mensagem, bem como o estudo do processamento e de técnicas aplicadas aos computadores e seus sistemas de programação. É uma ciência interdisciplinar derivada de campos relacionados, tais como a Matemática, Lógica, Linguística, Psicologia, Ciência da Computação, Engenharia da Produção, Artes Gráficas, Comunicação, Biblioteconomia, Administração, e outros campos científicos semelhantes.

Capurro, e Hjørland (2003) (2003), afirmam que a ciência da informação nasce em meados do século XX com um paradigma físico, questionado por um enfoque cognitivo idealista e individualista, sendo este por sua vez substituído por um paradigma pragmático e social ou, para tomar um famoso conceito: por uma "epistemologia social" (*social epistemology*), mas agora de corte tecnológico digital. Um número recente da revista *Social Epistemology* (v.16, n.1, 2002) é dedicado à relação entre epistemologia social e ciência da informação. Como se pode ver, o que aparentemente surge no final desse relativamente curto processo histórico, a saber, o paradigma social, já se encontrava no início, se bem que, não

como paradigma da ciência da informação, mas sim de seus predecessores, em particular a biblioteconomia e a documentação.

A Ciência da Informação, preocupada em esclarecer um problema social concreto, o da informação, voltada para o ser social que procura a informação, situa-se no campo das ciências sociais (das ciências do homem e da sociedade), que são o meio principal de acesso a uma compreensão do social e do cultural.

Tomando-se por base esta definição de Borko, ficam evidente e indiscutível que a Ciência da Informação tem característica interdisciplinar, ou seja, permeia diversos campos ou áreas do conhecimento.

A interdisciplinaridade traduz-se por uma colaboração entre diversas disciplinas que leva a interação, isto é, certa reciprocidade nas trocas, de modo que haja, em suma, enriquecimento mútuo. A forma mais simples de ligação é o isomorfismo, a analogia.

Borko (1968, p 2) ainda reforça que,

Ciência da Informação como uma disciplina tem como meta fornecer um corpus teórico sobre informação que propiciará a melhoria de várias instituições e procedimentos dedicados à acumulação e transmissão de conhecimento. Há um número significativo de instituições e meios de comunicação relacionados à área, e incluem: livros, visando o empacotamento do conhecimento; escolas para ensinar sobre as questões que envolvem o conhecimento acumulado de muitas gerações; bibliotecas para armazenar e disseminar conhecimento; filmes e televisão para a exposição visual de conhecimento; periódicos para a comunicação escrita dos últimos avanços técnicos em campos especializados; e conferências para as comunicações orais de informação.

Capurro, e Hjørland (2003) afirmam que o desenvolvimento da computação e a investigação empírica dos processos neuronais cerebrais vêm revolucionando a teoria clássica do conhecimento baseada na ideia da representação, ou duplicação de uma realidade externa na mente do observador. Essa revolução começa com a chamada teoria da informação de Claude Shannon e Warren Weaver (Shannon/Weaver 1972) e com a cibernética, que em meados do século passado tematiza o vínculo entre os seres vivos e, em geral, entre sistemas logo chamados autopoieticos e o meio ambiente como se expressou, claramente, Norbert Wiener (1961).

A Ciência da Informação é uma das novas interdisciplinas, um desses novos campos de conhecimento onde colaboram entre si, principalmente: psicologia, linguística, estatística,

eletrônica, economia, direito, filosofia, política e telecomunicações, nas palavras de Coadic (2004).

Trabalhar a exponencial massa de dados operacionais diariamente e convertê-la em informações que possibilitem a tomada de decisões e agreguem valor às atividades e aos negócios é um dos grandes desafios da Ciência da Informação. Nesse processo de transformação, tem ocorrido uma rearticulação das relações sociais e de produção em torno das TIC, gerando um deslocamento das instâncias de mediação política, econômica e social, da dimensão espacial para a temporal, e a construção e instalação do princípio de instantaneidade e de emergência como base de regulação de nossa experiência significativa. Um novo conceito de velocidade é instaurado, não mais físico, mas cognitivo, relativo não mais às distâncias físicas, mas sim ao volume de informação gerada ou acessada.

Assimilar os ambientes informacionais digitais, com sujeitos psicossociais autônomos conectados em rede, requer estudos interdisciplinares que suportem a complexidade e a qualidade informacional que os constitui. É também nesse sentido que as TIC vão além de ferramentas; é necessário considerar que o processo de globalização apresenta as relações de sentido em um novo quadro conceitual, marcado por avanços constantes tecnológicos e informacionais em uma perspectiva social.

Encontrar uma representação precisa e única é algo implícito no contexto, porém construir uma estrutura que contemple formas diferentes de apresentações com o propósito específico de ajustá-la de acordo com o problema a ser resolvido e de acordo com as habilidades cognitivas do usuário, e ainda garantir a interoperabilidade entre sistemas é outro desafio a ser considerado. O conhecimento se desenvolve a partir de uma interação entre quem busca o conhecimento e as formas de representação, ou notações disponíveis; fala-se de um conhecimento que era ação, ou seja, processos são dirigidos por uma forma de representação que guia as decisões de um indivíduo. O processo de obter conhecimento envolve diversas dimensões. Ajustar as circunstâncias, aos usuários, a ontologia e as tarefas envolvidas constituem essas dimensões, sendo esta última a mais importante.

Segundo Alvarenga (2003), os processos de representação são de dois tipos: a representação primária, que ocorre quando são criados documentos como objetos que representam o conhecimento, e a representação secundária, que visa à organização de

documentos com a inclusão em sistemas referenciais, ou seja, através da criação de metadados.

Conforme Souza et al. (2015), no tocante aos processos de representação da informação e do conhecimento, tão erroneamente ligados às origens da área, corre-se o risco de tornar-se a CI um espectador passivo de um rapidamente cambiante panorama. Apesar de se ter primazia no estudo e construção de instrumentos como as linguagens documentárias e interfaces de sistemas de informação, estes temas são hoje associados às ciências mais técnicas, que com mais propriedade os têm incorporados aos seus fazeres e construtos.

De acordo com Kobashi (2008), Inteligência Artificial tem também apresentado propostas para aprimorar as formas de organizar e representar informações. Sowa (2000) afirma que os conteúdos informacionais, para serem recuperados, devem ser categorizados por meio de atributos relacionados lógica e semanticamente. As ontologias são, nessa concepção, os sistemas de referência para categorizá-los e representá-los. Em uma ontologia, são descritas as categorias de coisas que existem ou podem existir em um domínio de aplicação. Para dar suporte ao raciocínio sobre as coisas de um domínio, a representação do conhecimento deve descrever o comportamento das coisas e as suas interações. As ontologias são, em larga medida, correlatos eletrônicos de tesouros, pois apresentam estrutura semelhante a estes últimos: termos de domínios específicos, suas definições e relações entre os termos.

Tendo agora, aspectos gerais sobre informação, conhecimento e dado, é preciso considerar a que domínio se aplica. A partir de então, pode-se pensar em outros contextos que são organização e representação do conhecimento.

Ainda, segundo Souza (2013), citado por Alvares (2012), observam-se os progressos realizados pela ciência da computação no escopo da Web Semântica e no campo específico dos *Knowledge Organization Systems* (W3C, 2009; NKOS, 2009); das tecnologias e metodologias para construção de ontologias (SILVA, 2008); nos sistemas de recuperação de informação (SOUZA, 2006) e nos processos de análise de domínios e modelização (WAND; STOREY; WEBER, 1999).

Há muito que o campo da ciência da computação deixou de ser considerado como aquele que provê tecnologias e sistemas; e tomou a frente de áreas nucleares de organização de informação e conhecimento, ensejando novos diálogos e repertórios de parte da Ciência da

Informação. Exemplos de áreas relativamente novas e pujantes não faltam, como as citadas, processamento de linguagem natural e a linguística computacional, além de uma recrudescida inteligência artificial, que vêm desenvolvendo pesquisas em classificação e indexação automáticas, construção automática de ontologias, redes neurais, análise do discurso, além de alimentar as pesquisas de ponta em *text mining*, *web mining*, *data mining* e assemelhados (SOUZA, 2013).

## 2.2. A ÁREA DE ORGANIZAÇÃO DO CONHECIMENTO

Organização do Conhecimento tem grande importância devido à atividade voltada para classificação do conhecimento em sistemas de recuperação de documentos. A exemplo dos tesouros, entre outros, seus esquemas de classificação foram criados para organização ou representação de conteúdos visando acesso, disseminação e recuperação.

De acordo com Souza (2006, p. 27)

A Organização do Conhecimento é área tradicional de pesquisa e ensino em Ciência da Informação. Grande parte da literatura nesta área trata do processo, das atividades e dos instrumentos especialmente desenvolvidos no tratamento de documentos para armazenamento, disseminação, recuperação e uso em sistemas e serviços de informação. Os diferentes tipos e demandas de informação têm demonstrado a crescente diversidade em organização do conhecimento em outros contextos de aplicação.

A representação do conhecimento é instrumento de utilização de várias áreas do conhecimento e, com isso, há uma variedade no que diz respeito a o que representar e como representa-lo. Entender o fundamento da representação, ou seja, seu significado e sua extensão passa a ser tarefa da Ciência da Informação.

Para Gomes (2011) organização do conhecimento é uma área que procura estabelecer suas bases teóricas, incluindo as diversas abordagens epistemológicas com suas implicações na elaboração de sistemas de organização do conhecimento, em que releva o papel da classificação. As pesquisas atuais apontam para as dificuldades daqueles sistemas face à multidisciplinaridade dos temas de pesquisa e às questões de classificação dos registros, seja em meio convencional, seja em meio digital. Sendo ela mesma uma área multidisciplinar, busca em outras áreas apoio para a organização dos sistemas.

De acordo com Souza (2006), tradicionalmente, as atividades de organização do conhecimento e representação da informação estiveram relacionadas a sistemas de

recuperação de documentos. Os esquemas de classificação, gerais e especializados, os tesouros, entre outros tipos de instrumentos, foram criados para a organização física de acervos ou para a representação temática do conteúdo intelectual dos documentos visando acesso, disseminação e recuperação sistemática.

Considerando a variedade de padrões, diretrizes e aplicações, nasce a ISKO em 1989 como a principal sociedade científica responsável pela área de Organização do Conhecimento.

Para alcançar seus objetivos, a ISKO<sup>7</sup> trabalha para:

- Promover a pesquisa, o desenvolvimento e aplicações de sistemas de organização conceituais do conhecimento que promovam o estudo dos aspectos filosóficos e semânticos da estrutura do conhecimento;
- Proporcionar os meios de comunicação e redes em organização do conhecimento para os seus membros; e,
- Funcionar como ponto de rede entre instituições e sociedades nacionais que trabalham com questões relacionadas à organização conceitual e à dinâmica do conhecimento.

Para a ISKO, a organização do conhecimento é considerada um foco disciplinar distinto da ciência da informação, com fortes conexões com outros domínios intelectuais, como filosofia, ciência da computação, psicologia, sociologia e muito mais.

Sendo assim, temos uma instituição que, especificamente trata a organização do conhecimento com escopo definido que segue parâmetro baseados em publicação artigos de pesquisa originais que:

1. Esclarecem os fundamentos teóricos (teoria de ordenação geral, fundamentos filosóficos do conhecimento e seus artefatos, bases teóricas de classificação, análise e redução de dados);
2. Descrever operações práticas associadas à indexação e classificação, bem como aplicações de sistemas de classificação e tesouros, indexação manual e de máquina;
3. Traçar a história da organização do conhecimento;
4. Discutir questões de educação e treinamento em classificação;

---

<sup>7</sup> ISKO Internacional Disponível em <[http://isko-brasil.org.br/?page\\_id=5](http://isko-brasil.org.br/?page_id=5)>. Acesso em 5 fev 2020

## 5. Problemas de terminologia em geral e com relação a campos especiais.

Para Esteban Navarro (1996) a Organização do Conhecimento é a disciplina da Ciência da Informação que se dedica ao estudo dos fundamentos teóricos do tratamento e da recuperação da informação, construção, manutenção, avaliação e uso dos instrumentos lógico-linguísticos mais adequados para controlar os processos de representação, classificação, ordenação e armazenamento do conteúdo informativo dos documentos com a finalidade de permitir sua recuperação e comunicação.

Dahlberg (2006), ao discutir a Organização do Conhecimento, declara que seu campo de assunto é formado por uma grande conjuntura conceitual e questiona se é possível falarmos de uma nova disciplina considerando-se o único e exclusivo campo de uma extensa massa de documentos publicados.

São muitas as áreas, como dito anteriormente que se aproximam da organização do conhecimento como, por exemplo, a gestão do conhecimento o que faz pensarmos que a organização do conhecimento tem várias aplicações ou encaixes.

Conforme Gomes (2011) além das discussões teórico-filosóficas e de novas formas de produzir o conhecimento, a Organização do Conhecimento se envolve com um lado de aplicação, o de produção de instrumentos. Aqui, também, existem teorias e métodos igualmente relevantes, voltados para os SOC. Desenvolver sistemas de classificação significa estabelecer o sistema de conceitos / termos a serem usados na indexação, na recuperação e no acesso / índice. Devem conter uma sintaxe, notação (no caso das tabelas), bem como a terminologia de acesso ao sistema e usada como base da indexação.

Com o advento do ambiente Web, novos instrumentos de representação organização e representação do conhecimento têm sido propostos bem como implementados, determinando novas discussões temáticas convergindo para diversos campos temáticos, repercutindo em atendimento de demandas informacionais de uma sociedade cada vez mais calcada em uma economia informacional de larga escala.

Organização do Conhecimento é um campo de pesquisa, ensino e prática, que é principalmente afiliado à biblioteconomia e à ciência da informação. A Organização do Conhecimento é, em primeiro lugar, institucionalizada em professores de universidades de todo o mundo, em programas de ensino e pesquisa em instituições de pesquisa e escolas de

ensino superior, em periódicos acadêmicos, em conferências nacionais e internacionais e em organizações nacionais e internacionais (DAHLBERG, 2010).

A Organização do Conhecimento descreve, representa, arquiva e organiza documentos e representações de documentos, bem como assuntos e conceitos tanto por humanos quanto por programas de computador (Hjørland, 2008). Para esses fins, são desenvolvidos regras e padrões, incluindo sistemas de classificação, listas de títulos de assuntos, tesouros e outras formas de metadados. A Organização do Conhecimento em termos de sistemas de classificação e sistemas de conceitos é um assunto central na organização do conhecimento. Os dois principais aspectos da organização do conhecimento são (1) processos de organização do conhecimento e (2) sistemas de organização do conhecimento. Os processos de organização do conhecimento são, por exemplo, os processos de catalogação, análise de assuntos, indexação, marcação e classificação por seres humanos ou computadores. Os sistemas de organização do conhecimento são a seleção de conceitos com uma indicação de relações semânticas selecionadas. Exemplos são sistemas de classificação, listas de títulos de assuntos, tesouros, ontologias e outros sistemas de metadados.

A Internet e seus mecanismos de pesquisa revolucionaram a maneira como as pessoas pesquisam e encontram informações. Quando comparado aos bancos de dados clássicos, que exigem especialistas em informações profissionais ou buscam usuários finais competentes, os mecanismos de pesquisa são (ou parecem ser) muito fáceis de usar. Além disso, os mecanismos de pesquisa têm uma cobertura ampla e abrangente de vários tipos de documentos. A Internet se tornou o meio mais importante para organizar e pesquisar informações e documentos. O campo conhecido como arquitetura da informação foi desenvolvido como um novo campo, que se preocupa em organizar o conhecimento sobre a WWW. Seu meio é novo, mas seus princípios básicos fazem parte do campo da Organização do Conhecimento, conforme definido no início deste artigo.

Paralelamente ao desenvolvimento da Internet, um novo tipo de sistema de organização do conhecimento, denominado "ontologia", tornou-se importante a partir dos anos 90. Dagobert Soergel escreveu:

A classificação há muito tempo é usada em bibliotecas e sistemas de informação para fornecer orientação ao usuário para esclarecer suas necessidades de informação e estruturar os resultados de pesquisa para navegação, funções amplamente ignoradas pela comunidade de recuperação de texto, mas agora recebendo atenção crescente no contexto de ajudar os usuários a lidar com os problemas da grande quantidade de informações na Web. Recentemente, outros campos, como

inteligência artificial, processamento de linguagem natural e engenharia de software, descobriram a necessidade de classificação, levando ao surgimento do que esses campos chamam de ontologias (Soergel 1999, 1119). Mas uma classificação por qualquer outro nome ainda é uma classificação (Soergel 1999, 1120).

Vickery (1992, p.195) observou apropriadamente que

a representação do conhecimento em forma simbólica é uma questão que ocupou o mundo da documentação desde sua origem. O problema agora é relevante em muitas situações além de documentos e índices. A estrutura de registros e arquivos em bancos de dados; estruturas de dados em programação de computadores; a estrutura sintática e semântica das línguas naturais; representação do conhecimento em Inteligência Artificial; modelos da memória humana; em tudo isso é necessário decidir como o conhecimento pode ser representado para que a representação possa ser manipulada.

A representação de conhecimento é a chave para a recuperação e disseminação efetiva de dados, informações e conhecimentos, conforme destaca Vickery (1992, p.196)

Ranganathan define conhecimento como a totalidade de ideias conservadas por seres humanos. Ele considera o conhecimento igual para o universo das ideias. Ao mesmo tempo, ele define informações como ideias comunicadas por outros ou obtidas por estudo e investigação pessoal. Um olhar mais atento em ambas as definições deixa claro que a o ponto convergência é Ideia. Também é evidente que Ranganathan usou "Informação" e "Conhecimento" como sinônimos. Ele o próprio diz que os termos "pensamento", "conhecimento" e "informações" são frequentemente usados como sinônimos do termo "Ideia". A ideia, segundo ele, é "o produto de pensando, refletindo, imaginando etc., conquistados pelo intelecto integrando com o auxílio da lógica, uma seleção da percepção de massa e / ou o que é diretamente apreendido por intuição e depositada na memória". Percepção de massa denota os conceitos já presentes em memória, à qual os recém-recebidos percebem e conceitos formados são assimilados. Os conceitos são a forma depositadas na memória, como resultado da associação de percepções - tanto puras quanto compostas - já depositadas na memória. As percepções são impressões significativas produzidos por entidades através dos sentidos primários e depositado na memória. O conhecimento evolui através da formação de percepções, conceitos e ideias. As percepções são experiências sensoriais, enquanto conceitos e ideias constituem experiências intelectuais. Em outras palavras, percepções são fenômenos sensoriais e perceptivos, enquanto conceitos e ideias são conceituais e, portanto, no nível cognitivo de percepção. Sendo assim, pode-se concluir que o conhecimento é o resultado de experiências sensoriais e intelectuais. Pode ser amplificado por experiências intuitivas, que são raras e ocasionais na natureza.

A organização da informação e a organização do conhecimento têm importância significativa para a Ciência da Informação. Há desafios ainda não superados, independentes dos avanços tecnológicos decorrentes de necessidades apontadas, ocorridos no final do século passado (ALVARES, 2014).

Com isso, sobrevêm questionamentos, notadamente de necessidades diárias, que vão de encontro a pressupostos desconhecimentos quanto à como analisar informação e conhecimento, como entendemos a representação prática e quais padrões estão à disposição para eventuais suportes de organização e recuperação da informação.

Há contraposições e sobreposições no árduo caminho de adequação da linguagem natural, da cognição humana e a complexidade dos processos notadamente particulares à natureza representativa seja da informação, seja do conhecimento ou ambos (LE COADIC, 2004).

Para Ramalho (2015) conceitos são invenções da mente humana usados para construir um modelo de mundo. Eles empacotam a realidade em unidades discretas para posterior processamento, suportam poderosos mecanismos lógicos, e são indispensáveis para cadeias de raciocínios extensas e precisas. Contudo, conceitos e percepções não podem formar um modelo perfeito de mundo, eles são abstrações que selecionam características importantes para determinado propósito, mas ignoram detalhes e complexidades que podem ser importantes para outros propósitos.

Ainda menciona Ramalho (2010) que o aumento exponencial no volume de recursos informacionais disponíveis, principalmente no meio digital, tem acarretado um crescente impacto no âmbito de atuação dos profissionais da informação, provocando verdadeiras mudanças paradigmáticas na área de Ciência da Informação e impondo o desenvolvimento de instrumentos e métodos que possibilitem suprir as novas necessidades informacionais.

De acordo com Alvares (2012) no sentido mais genérico do termo, organização do conhecimento é o modo como ele é disposto em assuntos em toda parte onde se deseja a sua sistematização ordenada para atingir determinado propósito. Pelo seu caráter interdisciplinar, a organização do conhecimento é estudada também em outras áreas, como antropologia, computação, filosofia, linguística, psicologia, sociologia, entre outros.

A Ciência da Informação, possui ferramentas e práticas voltadas à atividades de organização, representação e recuperação da informação tentando, dentro de seus limites, dar respostas a como representar o conhecimento, se as áreas de conhecimento são representadas da mesma forma, o que pode ser representado e se tudo pode ser representado (COADIC, 2004).

De acordo com Brascher e Café (2008), na área de Ciência da Informação, muito tem sido discutido sobre os termos informação e conhecimento. Esta sessão expõe características conceituais que distinguem os dois termos, com o objetivo de subsidiar nosso estudo sobre OI, OC, RI e RC. Sabemos das diferentes visões encontradas na área acerca de informação e

conhecimento, no entanto, não é objetivo de nosso trabalho contrapor as ideias dos diversos estudiosos desse tema.

Nessa direção Alvares (2012) aponta também que:

Representar o conhecimento é um esforço, nas diversas ciências, de materializar o que ocorre na mente humana e na dinâmica do conhecimento, cuja estrutura e funcionamento ainda significam enorme desafio para a comunidade científica. Para isso são usados diversos modelos de representação, que podem ser descritivos, matemáticos, visuais, informatizados, dentre outros, conforme as necessidades específicas de cada área para se aproximar dos processos cognitivos que envolvem a aprendizagem, o raciocínio, as relações que ocorrem entre os conhecimentos disponíveis na memória para gerar novos conhecimentos.)

Para um sistema de informação ter qualidade na recuperação do conhecimento, ele depende diretamente dos procedimentos e instrumentos envolvidos direta ou indiretamente no processo de organização da informação, ou seja, toda a padronização de organização é feita em tempo de construção do sistema (ALVARES, 2012).

O resultado esperado da organização da informação é a posterior recuperação da informação registrada em quaisquer formatos sejam textos, imagens, contexto áudio visual, internet, e demais formatos, objetivamente de forma estruturada (ALVARES, 2012).

Organização e representação da informação e organização e representação do conhecimento estão em constante interação dadas as suas intersecções o que, por vezes decorrem uso inadvertido de um termo pelo outro como se fossem sinônimos, até pela complexa interação e interdependência que transitam entre conhecimento e informação e demais disciplinas (ALVARES, 2012).

Conforme Souza (2013), no tocante aos processos de representação da informação e do conhecimento, tão intimamente ligados às origens da área, corre-se o risco de tornar-se a Ciência da Informação um espectador passivo de um rapidamente cambiante panorama. Apesar de se ter primazia no estudo e construção de instrumentos como as linguagens documentárias e interfaces de sistemas de informação, estes temas são hoje associados às ciências mais técnicas, que com mais propriedade os têm incorporados aos seus fazeres e construtos.

### 3. SISTEMAS DE ORGANIZAÇÃO DO CONHECIMENTO

A Ciência da Informação, pela abrangência de seu objeto de estudo, possui fortes laços com outras disciplinas. Pela extensão de domínio, aproxima-se de questões epistemológicas com a semiótica ou linguística, ou com questões mais técnicas com as engenharias e Ciência da Computação. Nesse sentido, no que se refere às questões da recuperação da informação, de maneira automatizada, tem respaldo no compartilhamento com a Ciência da Computação no desenvolvimento de ferramentas (ALVARES, 2014).

Considerando a figura 2, as diferenças são aparentes quando ao grau de complexidade devido às formas de relacionamento existentes em cada tipo de vocabulário.

Figura 2. Diferenças quando ao grau de complexidade de cada tipo de vocabulário

Lista	Anel de Sinônimos	Taxonomia	Tesouro	Ontologia
<ul style="list-style-type: none"> <li>• Controle de ambiguidade</li> </ul>	<ul style="list-style-type: none"> <li>• Controle de sinônimos</li> </ul>	<ul style="list-style-type: none"> <li>• Controle de ambiguidade</li> <li>• Controle de sinônimos</li> <li>• Relacionamento hierárquico</li> </ul>	<ul style="list-style-type: none"> <li>• Controle de ambiguidade</li> <li>• Controle de sinônimos</li> <li>• Relacionamento hierárquico</li> <li>• Relacionamento associativo</li> </ul>	<ul style="list-style-type: none"> <li>• Controle de ambiguidade</li> <li>• Controle de sinônimos</li> <li>• Relacionamento hierárquico</li> <li>• Relacionamento associativo</li> <li>• Axiomas</li> <li>• Regras de inferência</li> </ul>

Adaptado da ANSI/NISO Z39.19-2005 (R2010)

O objetivo final dos SOC é auxiliar a recuperação da informação, e esse fim orienta todo o processo de construção de seu arcabouço teórico e prático, segundo Alvares (2012). O mérito dos vocabulários controlados é assegurar que conceitos distintos sejam descritos por formas linguísticas únicas. Na existência de várias formas, elas devem ser normalizadas de maneira que a informação destinada ao usuário atenda à necessidade dele.

Classificação, segundo Alvares (2012), está relacionada diretamente com a organização do conhecimento, como forma de categorizar e organizar determinada área, definindo a abrangência do objeto de estudo. A classificação é utilizada em praticamente todas as disciplinas, mas com foco de estudo é tratada, sobretudo, nas disciplinas da Ciência da Informação.

### **3.1. VOCABULÁRIOS CONTROLADOS**

Nas palavras de Kobashi (2008) as representações, para serem funcionais, devem exibir atributos, estes últimos, obtidos por extração de informação textual. Supõe-se, portanto, que os sistemas inteligentes compreendem textos, selecionam informação e os representam, tarefas que, devemos admitir, não são triviais. Além de representar o conhecimento, um sistema inteligente deve codificá-lo em uma linguagem que possa ser processada computacionalmente. As representações são expressas sob a forma de proposições lógicas.

Kobashi (2008) afirma que, por meio de linguagens lógicas é possível representar apenas conhecimento declarativo. Contudo, segundo esse mesmo autor, o conhecimento do mundo não pode ser encapsulado em enunciados declarativos. Assim, no atual estágio de desenvolvimento dos recursos computacionais e da inteligência artificial, não parece ser possível pensar a automatização completa dos processos de representação da informação.

Em princípio, é necessário que possamos diferenciar dois elementos de conceitos muito próximos: língua e linguagem.

Para Saussure (1975) linguagem é um fenômeno: é o exercício de uma faculdade que existe no homem. A língua é um conjunto de formas concordantes que esse fenômeno assume numa coletividade de indivíduos e numa época determinada.

A língua é uma convenção social de uma comunidade, um fenômeno social estabelecido em determinado espaço de tempo. A linguagem, por sua vez, torna-se mais ampla, é a capacidade de representar conceitos independentes da forma, por isso, torna-se mais flexível nas representações, podendo ser por meio da língua ou de outro sistema de representação, nas palavras de Alvares (2012).

Sendo assim, língua refere-se aos idiomas, que são sistemas sociais convencionados por uma comunidade de usuários em constante evolução, já a linguagem a linguagem pode representar sistemas de representação de conceitos com formas diversas como sons, sinais, a exemplo, linguagens de programação de computadores ou mesmo documentárias (ALVARES, 2012).

No tocante à Ciência da Informação, representar o conteúdo de documentos com perda mínima de informação é seu objetivo. A organização e recuperação da informação faz parte de seu arcabouço. Para tanto, conforme Alvares (2012), as linguagens convencionadas que traduzem os conceitos-chave dos documentos para linguagens reconhecidas por sistemas de informação ou de bibliotecas, conhecidas como linguagens documentárias ou linguagens de indexação.

A linguagem documentária tem como objetivo, nas palavras de Alvares (2012), representar, de forma padronizada, o conhecimento, de domínio específico e estabelecer a ligação entre o usuário e o conteúdo do documento. Ela é constituída artificialmente com o propósito de sintetizar, por meio de vocabulário controlado, o conteúdo de documentos, e ser utilizada em sistemas de indexação, armazenamento e recuperação da informação, delimitando o vocabulário de um domínio de forma clara e precisa.

Vocabulário controlado, conforme Alvares (2012) pode ser sintetizado como um rol de termos apresentados explicitamente, no qual esses termos devem ser definidos de forma não ambígua e não redundante.

De acordo com a Z39.19-2005 (R2010) , observar no mínimo , duas regras a seguir:

- Se o termo é usado para diferentes conceitos, ele deve ser explicitamente qualificado para solucionar essa ambiguidade;
- Se vários termos são usados para o mesmo conceito, um dos termos é identificado como o preferido no vocabulário controlado e os outros termos são listados como sinônimos ou equivalentes.

Em síntese, vocabulários controlados são estruturados para viabilizar a visualização de tipos de relacionamento deferentes entre termos que eles contêm.

Ainda de acordo com a ANSI/NISO Z39.19-2005 (R2010), existem quatro diferentes tipos de vocabulários controlados determinados pelo grau de complexidade estrutural que compreendem. São eles: as listas, o anel de sinônimos, as taxonomias e os tesouros.

É interessante observar, conforme Alvares (2012), que a taxonomia é organizada a partir de um ponto de vista predeterminado. Isso quer dizer que a elaboração da taxonomia depende do propósito a que se destina ou do entendimento da realidade de que a elabora. Não há taxonomia certa ou errada, nem geral o suficiente para abranger todos os aspectos da realidade. Por exemplo, o cão da figura 6 pode estar na subclasse da classe animais domésticos ou em outra estrutura taxonômica.

Alvares (2012 apud Cavalcanti, 1978), define Tesouro como uma lista estruturada de termos relacionados, utilizada por profissionais da informação para descrever um documento com a especificidade desejada e permitir aos usuários a recuperação da informação pretendida.

São definições de um Tesouro:

- É uma ferramenta voltada para a organização e recuperação da informação;
- Atende aos usuários: ao indexador que estabelece e define os termos e seus relacionamentos que representam documentos e ao consultante que busca a informação auxiliado pela ferramenta;
- É um tipo de vocabulário controlado expresso em linguagem documentária que formaliza os termos simples ou compostos e os relacionamentos entre eles;
- A estrutura dos tesouros permite a organização lógica e semântica de um domínio;
- A mesma estrutura também permite que os significados dos termos sejam limitados de forma unívoca;
- Essa restrição aos significados dos termos visa neutralizar a sinonímia, a polissemia e outras características comuns a linguagem natural;
- Tesouro oferece uma representação da estrutura conceitual de um domínio do conhecimento.

### 3.2. ONTOLOGIAS E REDES SEMÂNTICAS

As ontologias, da forma com que as conhecemos na Ciência da Informação, nasceram na Ciência da Computação para atender à necessidade de inferência artificial ou mecânica de fatos ou, a atualmente conhecida inteligência artificial (ALVARES, 2014).

A Ciência da Informação, pela abrangência de seu objeto de estudo, possui fortes laços com outras disciplinas. Pela extensão de domínio, aproxima-se de questões epistemológicas com a semiótica ou linguística, ou com questões mais técnicas com as engenharias e Ciência da Computação. Nesse sentido, no que se refere às questões da recuperação da informação, de maneira automatizada, tem respaldo no compartilhamento com a Ciência da Computação no desenvolvimento de ferramentas (ALVARES, 2012).

Nessa perspectiva, a Ciência da Informação e a Ciência da Computação compartilham as teorias relacionadas à Ontologia. Alvares (2012, p. 57) afirma que para a ciência da informação, ontologia é instrumento para representar o conhecimento expresso por meio da informação registrada, objeto de estudo da ciência da informação. Essa informação é carregada de conceitos. Nesse contexto, os termos denotam conceitos que estão inter-relacionados em dado domínio. São exatamente esses termos que povoam uma ontologia da área.

É importante ressaltar que as ontologias, vieram de pesquisas da Ciência da Computação para atendimento da necessidade da inferência artificial ou mecânica sobre fatos, aparentemente sem relação, que explicita novos relacionamentos. Isto é, a descoberta se dá no âmbito do raciocínio da máquina, por isso o termo cunhado Inteligência Artificial. Talvez a característica que mais se destaca e diferencia as ontologias dos demais tipos de vocabulário controlado seja a automação de seus relacionamentos, axiomas e regras de inferência (ALVARES, 2012).

De fato, para a Ciência da Informação, ontologia é instrumento para representar conhecimento expresso por meio da informação registrada, objeto de estudo da Ciência da Informação. Essa informação é carregada de conceitos. Nesse contexto, os termos denotam conceitos que estão inter-relacionados em certo domínio. São estes termos que povoam uma ontologia da área (ALVARES, 2012).

É sabido que a manutenção das ontologias, suas contradições bem como sua implementação, representam alguns dos desafios e até obstáculos a serem transpostos, porém, tê-la como ferramenta para a organização da Informação, a decorrente normalização do conhecimento de uma ou mais áreas determina melhoria na comunicação, amplitude no compartilhamento do saber e efetiva organização do conhecimento.

A gestão do conhecimento no início da história humana foi amplamente moldada pela comunicação oral antes da invenção das línguas, que então permitia que o conhecimento humano fosse registrado e transmitido através das gerações. Uma das primeiras abordagens de Representação do Conhecimento baseada em computadores são as Redes Semânticas, que representam conhecimento na forma de nós e arcos interconectados, onde nos representam objetos, conceitos ou situações, e arestas representam as relações entre eles (YAN et al. , 2016).

De acordo com Perez et al. (2017), um elemento chave de um Grafo de Conhecimento é a ontologia que ele tem como seu esquema. Em poucas palavras, a ontologia permite descrever e representar todas as informações presentes dentro do Grafo de Conhecimento.

A organização ontológica suporta a publicação de informações organizacionais através de vários domínios. A ontologia fornece termos para apoiar a representação de:

- Estrutura organizacional que abrange:
  - A noção de uma organização;
  - Decomposição em suborganizações e unidades;
  - Propósito e classificação de organizações.
- Estrutura de relatórios que abrange:
  - Estrutura de membros e relatórios em uma organização;
  - Papéis, posts e relacionamento entre pessoas e organizações.
  - Informações locais, que incluem sites ou edifícios, locais dentro de sites;
  - Histórico organizacional, que inclui fusão e renomeação.

Quanto à origem das Redes Semânticas, alguns pesquisadores argumentam que as Redes Semânticas vieram dos Grafos existenciais de Charles S. Peirce, enquanto muitos deles prestam homenagem a Quillian, que foi o primeiro a introduzir as Redes Semânticas em seus modelos de memória semântica. Memória semântica refere-se ao conhecimento geral (fatos,

conceitos e relacionamento), como uma cadeira. É diferente de outro tipo de memória de longo prazo, isto é, memória episódica, que se relaciona com alguns eventos específicos, tais como mover uma cadeira. Depois de Quillian, muitas variantes de Redes Semânticas foram propostas (YAN et al. , 2016).

Em comparação com formalismos de representação e raciocínio formais do conhecimento, como lógicas de predicados, as Redes Semânticas são relativamente fáceis de usar e manter. Por outro lado, eles sofrem de algumas limitações. Por exemplo, não há sintaxe e semântica para a Rede Semântica de Quillian. Isso deixa espaço para os usuários terem suas próprias interpretações de construtores em Redes Semânticas. Essa abordagem pode ser vista como flexível para alguns, mas também é criticada por dificultar a integração de redes semânticas, preservando seu significado original. Além disso, as Redes Semânticas não permitem que os usuários definam o significado de rótulos em nós e arcos (YAN et al. , 2016).

### **3.3. GRAFOS DE CONHECIMENTO**

De acordo com Terra (2002), o conhecimento tornou-se um recurso econômico proeminente, sendo mais importante que a matéria prima; mais importante, muitas vezes, que o capital envolvido. Considerados produtos econômicos, a informação e o conhecimento são mais importantes que qualquer outro produto da era industrial. O mapa do conhecimento nos revela como o conhecimento e os ativos do conhecimento operam e se manifestam, e passa-se a entender por que gerenciar o capital intelectual deve ser a prioridade número 1 de uma organização.

É difícil encontrar um único setor, empresa ou organização de qualquer espécie que não tenha passado a fazer uso intensivo da informação, que não tenha se tornado dependente do conhecimento, como fonte de atração para consumidores e clientes, e da tecnologia da informação como instrumento gerencial.

Remetendo-nos ao paradigma sociotécnico, Christopher Freeman (1982) citado por Castells (1999):

Um paradigma econômico e tecnológico é um agrupamento de inovações técnicas, organizacionais e administrativas inter-relacionadas cujas vantagens devem ser descobertas não apenas em uma nova gama de produtos e sistemas, mas, mas também e, sobretudo na dinâmica da estrutura dos custos relativos de todos os possíveis insumos para a produção. Em cada novo paradigma, um insumo

especificou conjunto de insumos pode ser descrito como o “fator- chave” desse paradigma caracterizado pela queda dos custos relativos e pela disponibilidade universal. A mudança contemporânea de paradigma pode ser vista como uma transferência de uma tecnologia baseada principalmente em insumos mais baratos de energia para outra que se baseia predominantemente em insumos baratos de informação derivados do avanço da tecnologia em microeletrônica e telecomunicações (Castells (1999) apud Christopher Freeman (1982).

Vale ressaltar, de acordo com Rogers (2018), que o papel dos dados para os negócios, hoje, está mudando drasticamente. Muitas empresas que, durante anos, usaram dados como parte específica de suas operações, estão agora descobrindo uma revolução dos dados: os dados estão sendo fornecidos por novas fontes, estão sendo aplicados a novos problemas e estão se tornando importante vetor de inovação.

Ainda com Castells (1999), o conceito de paradigma tecnológico, elaborado por Carlota Perez, Christopher Freeman e Giovanni Dosi, com a adaptação da análise clássica feita por Kuhn, ajuda a organizar a essência da transformação tecnológica atual, à medida que ela interage com a economia e a sociedade. Em vez de apenas aperfeiçoar a definição de modo a incluir os processos sociais além da economia, penso que seria útil destacar os aspectos centrais do paradigma da tecnologia da informação para que sirvam de guia em nossa futura jornada pelos caminhos da transformação social.

Características da base material da sociedade da informação:

- Tecnologias para agir sobre a informação
- Penetrabilidade dos efeitos das novas tecnologias
- Lógica de redes
- Flexibilidade

Convergência de tecnologias específicas para um sistema altamente integrado

Para Yan et al. (2016), as informações na Internet são fragmentadas e apresentadas em diferentes fontes de dados, o que torna a coleta automática de conhecimento e entendimento formidáveis para máquinas e até para humanos.

Os Grafos de Conhecimento, segundo PEREZ et al. (2017), tornaram-se predominantes nos círculos acadêmicos e da indústria durante anos, por ser um dos conhecimentos mais eficientes e eficazes em abordagens de integração. Técnicas para construção do Grafo de Conhecimento podem extrair informações de estruturas, fontes de dados semiestruturadas ou mesmo não estruturados e, finalmente, integram as informações ao

conhecimento, representadas em um grafo. Além disso, o Grafo de Conhecimento é capaz de organizar informações de uma maneira fácil de manter, fácil de entender e maneira fácil de usar.

Na realidade os Grafos de Conhecimento a ganharam maior destaque quando o Google passou a utilizar-se deste termo para descrever diferentes aplicações de representação de conhecimento, porém, as organizações estão reinventando-se conforme novas tendências e tecnologias que as tornam exponenciais, ou seja, o nivelamento de operações ultrapassa fronteiras da antiga globalização, por assim dizer, trazendo agilidade, colaboração, poder de criação e co-criação, fomentando criações de novos modelos de produtos e negócios mais assertivos e de maior visibilidade.

Conceitualmente, o Grafo de Conhecimento utiliza preceitos da lógica fuzzy, cujo objetivo é modelar a capacidade dos seres humanos de raciocinar em um ambiente com imprecisão. Assim, a lógica fuzzy<sup>8</sup> bem como, a teoria dos conjuntos podem ser úteis para agregar e representar conhecimento, especialmente se forem aplicadas técnicas de criatividade a exemplos de mapas mentais ou mesmo brainstormings (PEREZ et al. 2017).

Um Grafo de Conhecimento é um grafo semântico que consiste em vértices (ou nós) e arestas. Os vértices representam conceitos ou entidades. Um conceito refere-se às categorias gerais de objetos, como cientista, carro etc. Uma entidade é uma entidade física é um objeto no mundo real, como uma pessoa (por exemplo, Micheal Jordan), um local (por exemplo, Nova York) e uma organização (por exemplo, Nações Unidas). As arestas representam os relacionamentos semânticos entre conceitos ou entidades. Ademais, o Grafo do Conhecimento concentra entidades e conceitos fragmentados, portanto parcialmente observados, que podem ser conectados para formar uma solução completa e estruturada para um repositório de conhecimento, facilitando o gerenciamento, recuperação, uso e compreensão das informações que ele contém (YAN et al., 2016).

A tarefa principal para uma engenharia do conhecimento é extrair as informações úteis dos dados e, em seguida, integrar essas informações convertendo-as em conhecimento

---

<sup>8</sup> A lógica fuzzy foi introduzida nos meios científicos em 1965 por Lofti Asker Zadeh, através da publicação do artigo Fuzzy Sets no Jornal Information and Control. Hoje ela é elemento fundamental em diversos sistemas, sendo considerada uma técnica de excelência no universo computacional. Possui também enorme aceitação na área de controle de processos. O conceito fuzzy pode ser entendido como uma situação em que não é possível responder simplesmente "sim" ou "não". Mesmo conhecendo as informações necessárias sobre a situação, dizer algo entre "sim" e "não", como "talvez" ou "quase", torna-se mais apropriado, conforme Rignel et. All (2011).

abrangente e organizado. O Grafo de Conhecimento fornece uma estrutura geral para representar conhecimento, baseado na mineração e análise de entidades e seus relacionamentos.

A representação do conhecimento baseada em grafos tem sido pesquisada por décadas e o termo Grafo de Conhecimento não constitui uma nova tecnologia. Segundo Ehrlinger; Wöß (2018), os Grafos de Conhecimento estão no foco de pesquisas desde 2012, resultando em uma ampla variedade de descrições publicadas e definições. Não há atualmente definições de um núcleo comum conforme demonstra a tabela 1, logo em seguida:

Quadro 1. Definições de Grafo de Conhecimento

Definition	Source
“A knowledge graph (i) mainly describes real world entities and their interrelations, organized in a graph, (ii) defines possible classes and relations of entities in a schema, (iii) allows for potentially interrelating arbitrary entities with each other and (iv) covers various topical domains.”	Paulheim [16]
“Knowledge graphs are large networks of entities, their semantic types, properties, and relationships between entities.”	Journal of Web Semantics [12]
“Knowledge graphs could be envisaged as a network of all kind things which are relevant to a specific domain or to an organization. They are not limited to abstract concepts and relations but can also contain instances of things like documents and datasets.”	Semantic Web Company [3]
“We define a Knowledge Graph as an RDF graph. An RDF graph consists of a set of RDF triples where each RDF triple $(s, p, o)$ is an ordered set of the following RDF terms: a subject $s \in U \cup B$ , a predicate $p \in U$ , and an object $U \cup B \cup L$ . An RDF term is either a URI $u \in U$ , a blank node $b \in B$ , or a literal $l \in L$ .”	Färber et al. [7]
“[...] systems exist, [...], which use a variety of techniques to extract new knowledge, in the form of facts, from the web. These facts are interrelated, and hence, recently this extracted knowledge has been referred to as a knowledge graph.”	Pujara et al. [17]

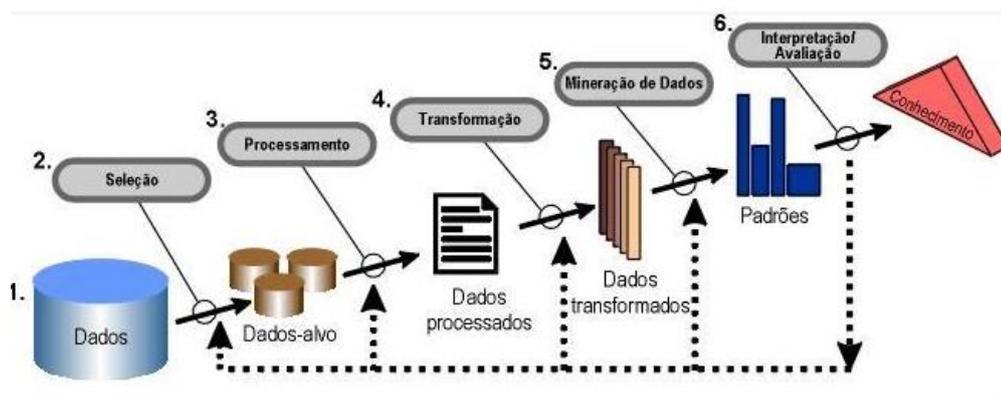
Adaptado de Ehrlinger e Wöß (2018)

Atualmente, as empresas estão estabelecidas em um ambiente caracterizado por quantidades crescentes de dados e, fazer agregação de conhecimento, representação e raciocínio tem sido altamente importante para o manuseio e atualização desses dados. Técnicas que envolvam criatividade são úteis e amigáveis para a aquisição do conhecimento. Como o raciocínio humano e seu ambiente são incertos e imprecisos, especialmente em contextos onde a criatividade é aplicada, os Grafo de Conhecimento são um formalismo útil para representar o conhecimento através de um compreensível caminho informacional.

Base de Conhecimento <sup>9</sup> é uma abordagem semelhante ao Grafo de Conhecimento, ou seja, refere-se a um banco de dados inteligente para gerenciar, armazenar e recuperar informações complexas estruturadas e não estruturadas. Em alguns casos, não há muitas distinções entre ambos, porém, a Base de Conhecimento enfatiza o conhecimento armazenado nas bases de dados, enquanto o Grafo de Conhecimento se concentra mais em estruturas gráficas. Podemos até admitir que um mecanismo sistemicamente e funcionalmente complementa o outro.

Para efeito ilustrativo, abaixo temos uma sequência de processos, conforme a figura 3, aludindo a estrutura de uma base de conhecimento com determinada finalidade.

Figura 3. Processo de descoberta de conhecimento em bases de dados.



Adaptado de Loureiro et al. (2014).

Conforme Yan (2016), o emprego de Grafos de Conhecimento é um exemplo razoável na era do Big Data, impulsionado pelos requisitos da indústria e motivações de pesquisa da academia. Atualmente, todos os principais mecanismos de pesquisa, como Google e Bing, empregam Grafos de Conhecimento para fornecer respostas dedicadas às consultas de pesquisa, como "esposa de Obama", onde a pergunta pode ser respondida diretamente por entidades, em vez de uma coleção de documentos relevantes da Web que contêm muitas informações redundantes.

<sup>9</sup> Indução de Árvores de Decisão na Descoberta de Conhecimento: Caso de Empresa de Organização de Eventos - Scientific Figure on ResearchGate. Available from: [https://www.researchgate.net/figure/Figura-3-Processo-de-descoberta-de-conhecimento-em-bases-de-dados\\_fig1\\_270449012](https://www.researchgate.net/figure/Figura-3-Processo-de-descoberta-de-conhecimento-em-bases-de-dados_fig1_270449012) [accessed 7 Sep, 2019].

De acordo com Yan et al. 2016, os Grafos de Conhecimento são capazes de inferir os significados conceituais das consultas e identificar as tarefas do usuário na pesquisa na Web, o que pode levar a obter melhores sugestões de consulta. O conhecimento concentrado no Grafo de Conhecimento permite que as máquinas entendam o texto em linguagem natural e construa tabelas da Web semiestruturadas.

Na Tabela 2, podemos observar que os tamanhos dos grafos de conhecimento se tornaram cada vez maiores e que recentemente, quase todas as empresas de Internet do mundo fornecem os serviços baseados nas técnicas de grafos de conhecimento. Estas mudanças se beneficiam da explosão no volume de dados e o progresso nas técnicas.

Tabela 2. Comparação entre Grafos de Conhecimento

Project	Affiliation	Year	KGs content	Data source
CYC	Cycorp	1984	120K+ concepts	Human-defined
WordNet(M)	PU	1985	117K synsets	Expert-authored
FrameNet(M)	UC Berkeley	1997	10 000 word senses and 170 000 annotated sentences	Human-defined
ConceptNet	MIT	1999	1.6M assertions	OMCS
HowNet	Keenage	2000	800 sememes	Modern Chinese dictionary
KnowItAll	UW	2005	54 753 facts	Web pages
TextRunner	UW	2007	500M assertions	117M+ Web pages
Freebase	Metaweb	2007	22M+ entities	Crowdsourced
DBpedia	DBpedia	2007	320 classes, 0.7M Wiki types, 3.6M entities, 247 M triples	Wikipedia
YAGO/YAGO2	MPII	2007	10M+ entities, 120M+ facts	Wikipedia, WordNet
Kylin/KOG	UW	2008	Classified documents and sentences	Wikipedia infoboxes and WordNet
OpenCalais	Calais	2008	Semantic metadata	Web Pages
Satori	MS	2009	400M entities	Bing search streams
NELL	CMU	2010	123 categories and 55 relations	500M Web pages
Probase	MSRA	2011	2.7M concepts	1.68B Web pages
EntityCube/Renlifang(C)	MSRA	2011	Entity-Relation Graph	Billions of public Web pages
Google Knowledge Graph(M)	Google	2012	570M+ objects, 18B facts	CIA World Factbook, Freebase, Wikipedia
Sogou Zhilifang(C)	Sogou	2012	100M+ entities, 1M+ relations	Web pages
Baidu Zhixin(C)	Baidu	2012	N/A	Web pages
Facebook Graph Search	Facebook	2013	User's network of friends, Bing's search engine	700TB post and comment data from Facebook blog

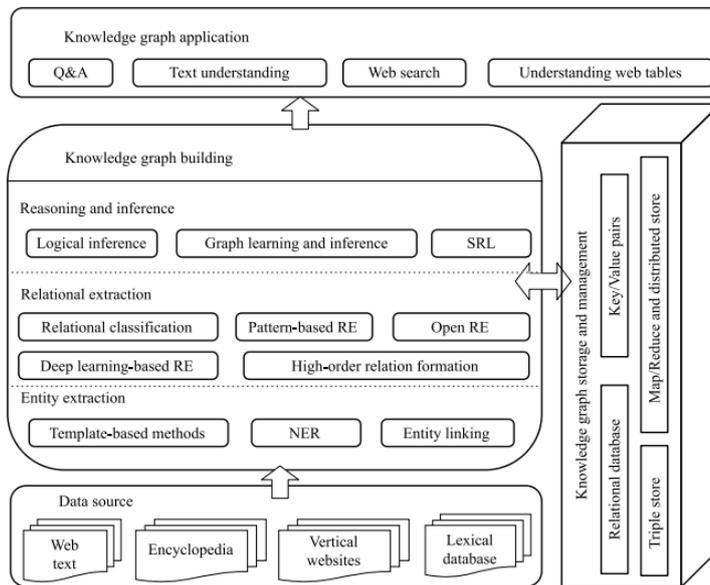
Adaptado de YAN et al.(2014)

A construção bem-sucedida de um Grafo de Conhecimento trará funcionalidade inteligente em várias aplicações. Com a existência de um Grafo de Conhecimento de alta qualidade e granulação fina, o mecanismo de pesquisa pode entender a consulta como “esposa de Obama” e recupere imediatamente a resposta simplesmente reconhecendo “Obama” como a menção em texto para a entidade “Barack Obama” e procurando o nome da entidade que tem o relacionamento “é esposa de” com “Barack Obama”. Além de aplicativos relacionados

à pesquisa na Web, o Grafo de Conhecimento pode servir como fonte de dados para a construção de um banco de dados semântico que pode ser usado automaticamente por computadores, como uma Wikipédia online com dados estruturados em um formato estrito. Permite que os sistemas inteligentes de atendimento de perguntas respondam a perguntas levantadas por clientes humanos/computadores em idiomas semelhantes a consultas.

Ainda conforme Yan et al. , 2016, os Grafos de Conhecimento podem ser um repositório de conhecimento estruturado que suporta um vasto número de aplicativos relacionados à análise de Big Data. Em empresas de comércio eletrônico como Walmart, pesquisa e recomendação de produtos podem ser apoiados por Grafos de Conhecimento que contenham informações de produtos, vendedores, fabricantes, etc. O desenvolvimento de medicamentos pode ser beneficiado por Grafos de Conhecimento, onde redes de doenças em larga escala podem ajudar melhorar o diagnóstico médico. A Figura 4 ilustra a estrutura de um Grafo de Conhecimento obtendo uma visão geral que nos ajuda a entender a relação entre cada componente e seu todo.

Figura 4. Estrutura de um Grafo de Conhecimento.



Adaptado de Yan (2016)

Comparado a outros sistemas de informação orientados ao conhecimento, os recursos distintos dos Grafos de Conhecimento estão em sua combinação especial de estruturas de representação de conhecimento, processos de gerenciamento de informações e algoritmos de

busca. Posteriormente, os pesquisadores em Representação e Raciocínio do Conhecimento abordaram alguns problemas conhecidos na Rede Semântica ao padronizar a versão moderna da Rede Semântica, ou RDF (Resource Description Frameworks). Acontece que técnicas de representação de conhecimento, como o Grafo de Conhecimento ou Rede Semântica, são úteis não apenas para pesquisa na Web, mas também em muitos outros sistemas e aplicativos, incluindo gerenciamento de informações corporativas (PEREZ et al., 2017).

O RDF (Resource Description Framework) é um padrão moderno do W3C, abordando alguns dos problemas relacionados às redes semânticas clássicas em termos de falta de sintaxe formal e semântica. Por exemplo, a relação “é um” pode ser representada pela propriedade *subClassOf* no RDF, cuja semântica é claramente definida nas especificações do RDF. Deve-se ressaltar que o RDF não trata de todas as limitações de uma rede semântica, por exemplo, o RDF também não permite que os usuários definam conceitos. Isto é, no entanto, abordado pelo OWL (*Web Ontology Language*), um padrão W3C para definir vocabulários para Grafos RDF (YAN et al., 2016).

No OWL, a parte da relação não é uma relação interna como a propriedade *subClassOf*. Em vez disso, é uma relação definida pelo usuário que pode ser expressa usando o construtor existencial. As lógicas de descrição são a base do padrão OWL na Web Semântica. Com base no RDF e no OWL, o *Linked Data* (dados vinculados) é uma estrutura comum para publicar e compartilhar dados entre diferentes aplicativos e domínios, onde o RDF fornece um modelo de dados baseado em Grafo para descrever objetos. O OWL oferece uma maneira padrão de definir vocabulários para anotações de dados (YAN et al., 2016).

No paradigma de dados vinculados, os grafos RDF podem ser vinculados por meio de mapeamentos, incluindo mapeamento em nível de esquema (*subClassOf*) e mapeamento em nível de objeto. O foco do presente trabalho, portanto, é sobre a construção, compreensão e exploração de Grafo de Conhecimento em muitas organizações.

Entre outros recursos, o mais típico da perspectiva do usuário é que, em além de uma lista classificada de páginas da Web resultantes da pesquisa por palavras-chave, o Google também mostra um cartão de conhecimento estruturado à direita, que é uma pequena caixa que contém um trecho de informações resumidas sobre a entidade que provavelmente resolve a pesquisa. Esse cartão de conhecimento contém informações adicionais relevantes para a pesquisa, contribuindo para aliviar a carga do lado do usuário na busca de páginas da Web

relevantes para encontrar respostas manualmente. Além disso, sugerem-se relações com outras entidades no Grafo de Conhecimento, aumentando a sensação de acaso e estimulando uma maior exploração pelo usuário. Na maioria dos casos, esses cartões de conhecimento atendem suficientemente às necessidades de informações dos pesquisadores, melhorando significativamente a eficiência dos sistemas de pesquisa na Web, tanto em termos de tempo gasto por pesquisa quanto na qualidade dos resultados. Inspirados na história de sucesso do Google, os Grafos de Conhecimento estão ganhando impulso na arena da World Wide Web.

Nos últimos anos, testemunhamos uma crescente adoção industrial por outros gigantes da Internet, que incluem o Graph Search do Facebook e o Satori da Microsoft, esforços contínuos feitos em pesquisas industriais, por exemplo, Cofre do Conhecimento, postando eventos orientados para a comunidade (Tutorial de Grafo de Conhecimento em WWW20152; KG20143), participando de colaborações entre a academia e a indústria e o estabelecimento de empresas especializadas em áreas como Diffbot4 e Syapse. Todas essas iniciativas foram tomadas em ambientes acadêmicos e industriais, desenvolveram e estenderam ainda mais o conceito inicial do Grafo de Conhecimento, popularizado pelo Google.

Recursos adicionais, novos insights e várias aplicações foram introduzidos e, como consequência, a noção de Grafo de Conhecimento se tornou um termo muito mais amplo que engloba toda uma linha de esforços da comunidade por direito próprio, novos métodos e tecnologias. Para explicar melhor as diferenças sutis entre o Grafo de Conhecimento e os Dados vinculados, primeiro precisamos introduzir alguns conceitos básicos.

Ao aludirmos sua construção, um Grafo de Conhecimento é um conjunto de entidades determinadas (com atributos) que se interagem por relacionamentos determinados. Os tipos de entidades e relacionamentos são definidos em esquemas que são chamados de Ontologias. Entidade presente nos contextos da Ciência da Informação, o padrão RDF é utilizado para representar Grafos de Conhecimento, e outros dois padrões definem as ontologias (RDFS e OWL) e o padrão para consulta de Grafos de conhecimento (SOUZA; MARTINS; RAMALHO, 2018).

Um dos principais passos durante a criação de um Grafo de Conhecimento é decidir qual o conhecimento que você deseja capturar em seu Grafo. Isso faz parte da modelagem de dados atividade. Nesta fase você pode decidir usar um vocabulário leve e existente, mas você também pode decidir usar um esquema mais pesado. Com esquemas leves você nomeia

apenas as principais relações entre as entidades em seus dados (você nem sequer precisa nomear os tipos de entidades). Um esquema mais pesado conterá definições com restrições sobre o que tipos de entidade específicos significam em seu contexto.

Esquemas pesados podem ser usados por alguns algoritmos para automatizar (parcial ou totalmente) algumas tarefas mais tarde (como interconexão de dados ou anotação de texto e desambiguação).

Conforme Perez et al. (2017), se você decidir criar definições formais para suas entidades, poderá reutilizar várias décadas de experiência em desenvolvimento de ontologias (o termo técnico para criar definição formal de um domínio).

Conforme Souza (2015), de fato, ontologias são o meio mais utilizado mundialmente para estruturar informação e conhecimento, em campos de pesquisa relevantes para a sociedade humana, como a medicina, o direito, a engenharia, dentre outros. Ao contrário do que se imagina, a pesquisa em ontologias apresenta muitas similaridades reais com pesquisas consagradas em ciência da informação (AMEIDA, 2013), mesmo que muitas vezes o termo “ontologia” não seja adotado.

Ainda, seguindo Souza (2015), questões enfrentadas por cientistas da computação para a construção de vocabulários legíveis por máquina são questões com as quais os cientistas da informação já têm se deparado por muitos anos (LVICKERY, 1997). Existem similaridades e diferenças entre ontologias da computação e os tesouros e taxonomias, dois outros termos amplamente usados na ciência da informação. A possibilidade de restringir a linguagem natural parece ser o ponto de contato entre esses três tipos de estruturas.

Sua popularidade, conforme ressalta Ehrlinger e Wöß (2016), é claramente influenciada pela introdução do Grafo de Conhecimento do Google em 2012, e desde então o termo tem sido amplamente usado sem uma definição unívoca. Uma volumosa variedade de interpretações tem dificultado a evolução de um entendimento comum dos Grafo de Conhecimento. Numerosos trabalhos de pesquisa se referem ao Grafo de Conhecimento do Google, embora não haja documentação oficial sobre os métodos usados.

A agregação, representação e raciocínio do conhecimento são tarefas difíceis para muitos usuários da informação, o que requer frequentemente especialistas dedicados a este respeito. O raciocínio humano ocorre em um ambiente impreciso e incerto, especialmente se a

criatividade estiver envolvida. Além disso, a linguagem natural, que é a base da comunicação humana e do raciocínio humano, é imprecisa e subjetiva.

O conhecimento agregado deve ser representado de forma a permitir que os sistemas de informação processem ativamente o conhecimento e, assim, possibilitar o raciocínio baseado no conhecimento. Assim sendo, os Grafos de conhecimento podem ser facilmente entendidos pelos usuários, e os conceitos subjacentes podem ser usados diretamente para o raciocínio desejado tornando-os uma ferramenta poderosa para a Representação do Conhecimento. Eles podem representar conhecimento explicitamente em uma base lógica e de forma estruturada. Assim, os usuários podem entender como a base de conhecimento está sendo construída, como é usada e ter o controle sobre cada etapa da construção dessa base de conhecimento.

A criação de conhecimento e gerenciamento de informações são questões prioritárias na mente dos gerentes como fontes potenciais de maior competitividade. Dentro dessa agenda mais ampla se concentram na tecnologia da informação e no grau que ela pode transformar a competitividade.

Primeiramente, as organizações podem fornecer esquemas unificados de conhecimento (ontologias) para seus negócios e espalhá-los ao longo dos ramos corporativos. Esses esquemas podem ser construídos com base em vocabulários de negócios e modelos de mercado existentes, e são facilmente viáveis para acompanhar a evolução do negócio. Geralmente apoiado por RDF e Grafos de armazenamentos de dados, os Grafos de Conhecimento provem melhor capacidade de evolução do esquema em relação às plataformas relacionais.

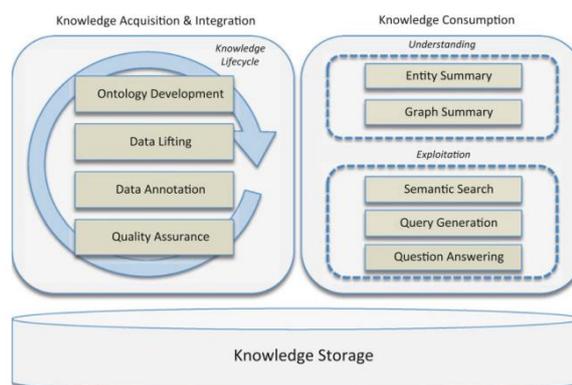
Isso facilita a adoção de modelos conceituais corporativos compartilhados, evitando assim a proliferação de esquemas heterogêneos para as mesmas entidades de negócios, uma das principais causas de ineficiência do gerenciamento de informações dentro de uma organização de negócios, conforme Perez et al. (2017).

No que diz respeito às entidades empresariais e consumidores singulares na Web, as empresas podem alavancar processos organizacionais estruturados e trabalhar com coleções controladas de fontes de informação, como em sistemas de *data warehouse*. Como nos processos de gestão da informação, na empresa tradicional, os Grafos de Conhecimento requerem a consolidação dos conjuntos de dados produzidos independentemente em

diferentes filiais e unidades de negócios, em um repositório centralizado para consumo subsequente.

Em termos de arquitetura (Figura 5), de acordo com Perez et al. (2017), há três tarefas principais relacionadas ao uso de Grafos de Conhecimento: construção, armazenamento e consumo.

Figura 5. Arquitetura para Grafo de Conhecimento em organizações



Adaptado de Perez et al. (2017).

Assim sendo, a arquitetura de um Grafo de Conhecimento se estrutura em:

**Aquisição de Conhecimento e Camada de Integração:** criação do Grafo de Conhecimento. Feito a partir dos dados já existentes para a extração do conhecimento em termos de entidades e suas relações - analisando os vários casos de uso e definindo conhecimento.

**Camada de Armazenamento do conhecimento:** armazenar o Grafo de Conhecimento de tal forma que será possível adicionar novos tipos de conhecimento e acessar o conhecimento codificado no Grafo de uma maneira eficiente.

**Camada de Consumo de Conhecimento:** o conhecimento codificado no Grafo de Conhecimento para uso da organização, a fim de melhorar a eficiência da organização.

A unidade básica de um Grafo de Conhecimento é (a representação de) uma entidade singular, como uma partida de futebol que você está assistindo, uma cidade que você visitará em breve ou qualquer coisa que você gostaria de descrever. Cada entidade pode ter vários atributos. Por exemplo, os atributos de uma pessoa incluem nome, data de nascimento,

nacionalidade etc. Além disso, as entidades são conectadas umas às outras por relações; por exemplo, você segue um de seus colegas no Twitter.

As relações podem ser usadas para conectar dois Grafos de Conhecimento separados. Por exemplo, dizendo que seu Identificador Unico do Twitter e o ID da sua carteira de motorista estão indicando a mesma pessoa, isso na verdade interliga os dados do Twitter com o espaço de informações na agência de licenciamento de motorista do seu país. Não é de surpreender que cada entidade precise de uma identificação para se distinguir. Este é o quebra-cabeça final na representação do conhecimento de um Grafo de Conhecimento. Observe que, para facilitar a interligação entre vários Grafo de Conhecimento, os IDs da entidade precisam ser globalmente exclusivos. Tipos de entidades e relações são definidos em alguns dicionários compreensíveis por máquina chamados ontologias.

A linguagem ontológica padrão é chamada OWL (Web Ontology Language). A qualidade de um Grafo de Conhecimento é crucial para suas aplicações. Por exemplo, um Grafo de Conhecimento deve ser consistente. No exemplo acima, pode ser que o seu endereço de contato na sua carteira de motorista seja diferente daquele no seu perfil do Twitter. Para criar um Grafo de Conhecimento conectando esses dois espaços de informações, essa inconsistência deve ser resolvida mantendo o correto. Além da consistência, é preciso também considerar a correção e a cobertura dos Grafo de Conhecimento, bem como a eficiência, tolerância a falhas e escalabilidade dos serviços com base em Grafo de Conhecimento. Muitos desses aspectos estão relacionados, entre outros, ao esquema (ontologia) de um Grafo de Conhecimento.

#### **4. PERSPECTIVAS DE UTILIZAÇÃO DE GRAFOS DO CONHECIMENTO EM AMBIENTES CORPORATIVOS**

De acordo com Perez et al. (2017), um sistema de informação baseado em Grafo de Conhecimento geralmente forma um ecossistema composto por três componentes principais: construção, armazenamento e consumo. As tecnologias relevantes de Grafo de Conhecimento podem ser classificadas em um desses componentes de um ecossistema em que sua contribuição é mais crítica. No que diz respeito à construção e armazenamento de Grafo de Conhecimento, encontram-se tecnologias e ferramentas para:

- Representação e raciocínio do conhecimento (idiomas, esquema e vocabulários padrão);
- Armazenamento de conhecimento (bancos de dados Grafo s e repositórios);
- Engenharia do conhecimento (metodologias, editores e padrões de design);
- Aprendizado (automático) do conhecimento, incluindo aprendizado de esquema e população.

Nos três primeiros itens, a maioria das tecnologias deriva das áreas de representação do conhecimento, Banco de Dados, Ontologias e Web Semântica. Por outro lado, para o aprendizado do conhecimento, normalmente são empregadas estruturas e tecnologias de Data Mining, Processamento de Linguagem Natural e Machine Learning. Do ponto de vista do consumo, o conteúdo dos Grafo de Conhecimento pode ser acessado e analisado diretamente por meio de linguagens de consulta, mecanismos de pesquisa, interfaces especializadas e / ou geração de resumos de Grafo s (domínio / aplicativo) específicos e análise visual. Em muitos outros casos, um Grafo de Conhecimento pode aumentar a eficácia de uma tarefa tradicional de processamento / acesso a informações (por exemplo, extração de informações, pesquisa, recomendação, resposta a perguntas etc.), fornecendo um valioso domínio em segundo plano conhecimento.

Conforme Perez et al. (2017), a comunicação entre marcas e consumidores está prestes a explodir. As características do produto não são mais a chave para as vendas e a combinação de recursos pessoais e coletivos benefícios está se tornando um aspecto cada vez mais crucial. De fato, as marcas que fornecem esse valor obtêm um impacto maior e, conseqüentemente, obtêm benefícios. Por outro lado, as novas tendências estão assumindo; induzindo uma mudança dramática na forma como consumidores e marcas se envolvem e quais canais e

tecnologias são necessários para ativar o processo. Como resultado, as fronteiras tradicionais dentro indústria estão sendo ampliadas e novas ideias, invenções e tecnologias são necessárias para acompanhar os desafios levantados pelas crescentes demandas desse uso intensivo de dados.

Assim, conforme Perez et al. (2017), é necessário alavancar avanços na área, estimulando um ecossistema de colaboração entre os diferentes atores. Exemplos inspiradores incluem a adoção pela Tesla Motors da política de patentes abertas, segundo a qual ela compartilha sua inovação no que diz respeito aos carros elétricos abertamente via Internet. Em troca, a Tesla espera que a indústria de automóveis evolua ainda mais o carro elétrico e dinamize o mercado. Na indústria de mídia, um caso paradigmático desta abordagem colaborativa é, Perez et al. (2017), o HAVAS 18 Innovation Labs, implantada em locais estratégicos em todo o mundo. Um desses locais é o Siliwood, centro de pesquisa em Santa Monica, co-criado com a Orange, que se concentra na convergência entre tecnologia, ciência de dados e conteúdo e mídia. O HAVAS 18 Innovation Labs procura conectar um mix de talentos locais nos sites, envolvendo inovadores, universidades, start-ups e tendências tecnológicas para iniciativas de co-criação relevantes, tanto para o HAVAS quanto para seus clientes, estarem um passo à frente. Para isso, o HAVAS criou um Grafo de Conhecimento corporativo e uma plataforma de informação que agrega todo o conhecimento disponível sobre tecnologia. Start-ups em todo o mundo tornam-se vivíveis para estrategistas de negócios de mídia através de um único ponto de entrada.

Até onde se conhece, o HAVAS esta é uma das primeiras aplicações dos princípios do Grafo de Conhecimento no mundo empresarial, e o primeiro na indústria de mídia, depois de gigantes de busca da Internet Google, Yahoo! e Bing, cada um com suas próprias implementações. Iniciativas relacionadas incluem esforços específicos de domínio, como Grafo s sociais do Facebook, e recursos de referência como DBPedia e Freebase, de acordo com Perez et al. (2017).

De acordo com Perez et al. (2017), outro exemplo que se segue, é o IBM Watson<sup>10</sup>, uma plataforma de computação cognitiva disponível na nuvem, desenvolvida pela IBM como

---

<sup>10</sup> Watson é a plataforma de serviços cognitivos da IBM para negócios. A cognição consiste no processo que a mente humana utiliza para adquirir conhecimento a partir de informações recebidas. Com o avanço da tecnologia, essa capacidade passa a ser integrada a sistemas que podem aprender em larga escala e ajudar a sociedade em uma série de finalidades, desde o atendimento a clientes até o combate a doenças graves, essa solução também é chamada de inteligência artificial.

um resultado do projeto “Jeopardy ! Desafio de perguntas e respostas”<sup>11</sup>. Watson usa processamento de linguagem natural e de máquina para aprender a descobrir insights de grandes quantidades de dados não estruturados, e fornece uma variedade de serviços para trabalhar com esse conhecimento. Grafos de Conhecimento (como Prismatic, DBPedia e YAGO) estavam no centro do sistema de perguntas e respostas da IBM.

Os serviços da Watson disponíveis atualmente oferecem recursos de Grafos de Conhecimento por meio de muitos serviços e interfaces de programas de aplicativos (APIs)<sup>12</sup>, como a expansão do conceito Watson. A pesquisa e o desenvolvimento em andamento visam estender a disponibilidade de bases de conhecimento estruturadas para os Serviços de Diálogo (*chatbots*) e outros *front ends* cognitivos.

A pesquisa tradicional e semântica (baseada em palavras-chave) são atualmente a principal maneira que os usuários do conhecimento usam para acessar informações organizacionais. Contudo, o paradigma de busca abrange apenas parte do processo que os usuários realizam. Normalmente, o usuário terá uma questão específica em mente, que pode ser traduzida como uma lista de palavras-chave para sua pesquisa. A principal desvantagem desta abordagem é que ela retorna uma lista de documentos que correspondem às palavras-chave pesquisadas. Depois da consulta, o usuário ainda precisará abrir os documentos, extrair a informação que eles estavam procurando (se está de fato nos documentos encontrados).

---

<sup>11</sup> Jeopardy! é um programa de televisão atualmente exibido pela CBS Television Distribution. É um show de perguntas e respostas variando história, literatura, cultura e ciências.

<sup>12</sup> API é um conjunto de rotinas e padrões de programação para acesso a um aplicativo de software ou plataforma baseado na Web. A sigla API refere-se ao termo em inglês "Application Programming Interface" que significa em tradução para o português "Interface de Programação de Aplicativos"

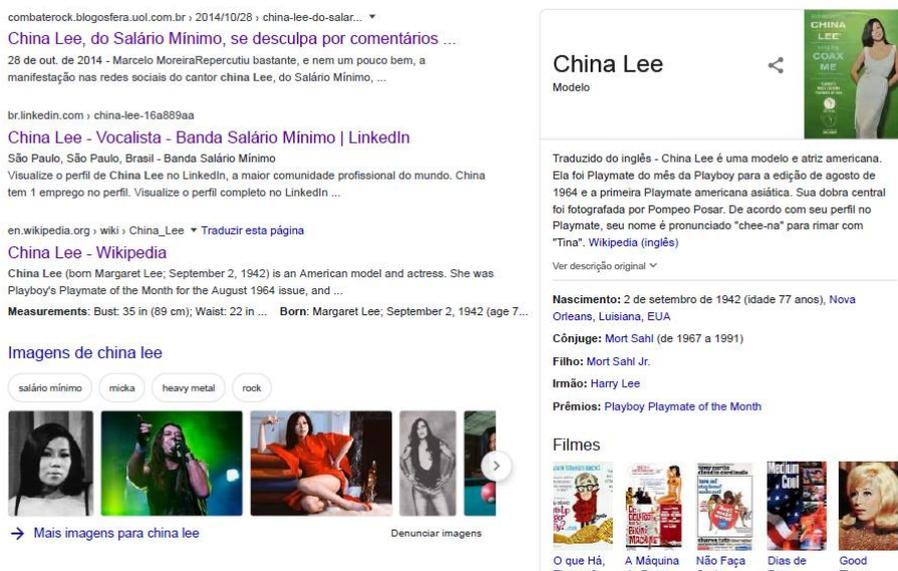
De acordo com Perez et al. (2017), o *Question Answering* visa automatizar este processo sempre que possível. A ideia principal é que, em vez de traduzir a questão em palavras-chave e, em seguida, lendo os documentos encontrados, o usuário pode perguntar diretamente e obter uma resposta, se a informação estiver disponível (Figura 6).

Figura 6. Exemplos de perguntas respondidas pelo Google Search Engine.



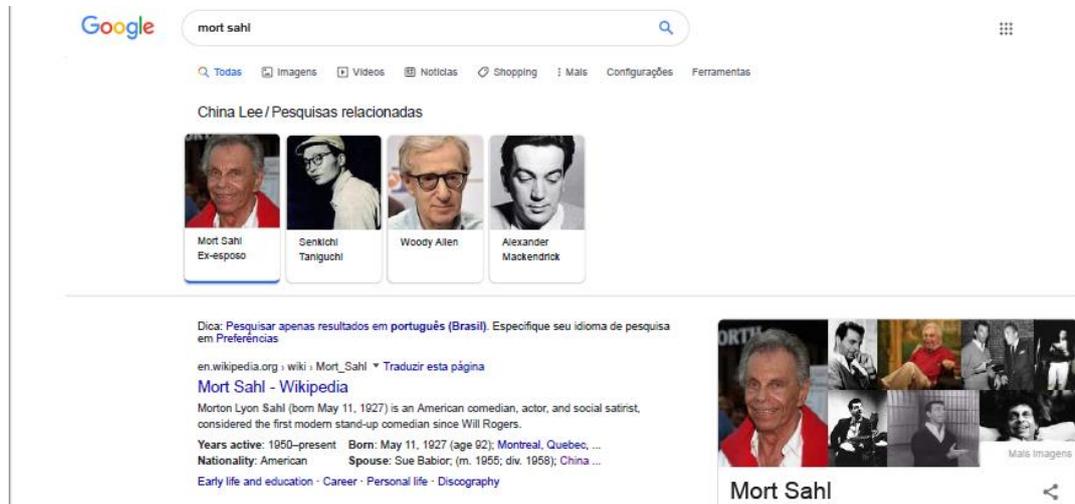
Adaptado de Perez et al. (2017)

Figura 6.1 Exemplos de perguntas respondidas pelo Google Search Engine.



Fonte: Google Search

Figura 6.2. Exemplos de perguntas respondidas pelo Google Search Engine.



Fonte: Google Search

Para que a pergunta seja respondida, o sistema deve ser capaz de analisar novas questões de linguagem natural (escritas ou faladas) e mapeá-las corretamente em consultas ao Grafo de Conhecimento (ou um mecanismo de busca semântico). Dependendo da pergunta e se a informação está disponível diretamente do Grafo de Conhecimento ou como parte de um documento previamente analisado, o sistema precisa determinar se pode responder a pergunta e como apresentar a resposta ao usuário.

Em 2008, conforme Perez et al. (2017) previa-se que as tendências atuais e futuras das tecnologias semânticas estivessem na interseção de três dimensões principais:

- Interação natural,
- a Web 2.0,
- Arquiteturas orientadas a serviços

Se nos afastarmos desses termos particulares, o significado real se tornará bastante simples:

- Facilidade de acesso aos sistemas de computadores pelos usuários finais,
- Capacitação das comunidades de usuários para representar, gerenciar e compartilhar conhecimento de maneiras colaborativas,
- Interoperabilidade da máquina.

Desde então, inúmeros desafios de pesquisa foram enfrentados em áreas como aquisição, representação e descoberta de conhecimento, metodologias de engenharia do conhecimento, vocabulários, arquiteturas escaláveis de gerenciamento de dados, interação homem-computador, recuperação de informações e inteligência artificial, onde tecnologias semânticas estão envolvidas, contribuindo para avanços cruciais em sistemas intensivos em conhecimento, o valor dos dados como força motriz por trás de aplicativos inteligentes permanece. No entanto, existe uma nova tendência que ganha impulso, que consiste na percepção de que esse valor é diretamente proporcional à interligação dos dados, não apenas em sistemas complexos e abertos como a Web, mas também em aplicativos empresariais específicos baseados em combinações de dados corporativos e abertos. Mais adequados para operações de pesquisa e relacionamento, dados mal formalizados, mas altamente interconectados estão se tornando mais populares do que conjuntos de dados isolados e formalizados.

Conforme Perez et al. (2017), o cenário atual de aplicativos, mais orientado para dispositivos móveis e em tempo real, está impondo essa nova mudança de paradigma. O Google entendeu isso muito bem e, em 2012, começou a impulsionar essa tendência no setor, liberando seu Grafo de Conhecimento como uma maneira de dominar esse valor, uma base de conhecimento que aprimora os resultados de seu mecanismo de pesquisa com informações de pesquisa semântica coletadas de uma ampla variedade de fontes. Curiosamente, o Grafo de Conhecimento fornece uma maneira de conectar os pontos (entidades) por meio de relações explícitas, com as entidades e relações descritas seguindo modelos formais (mas leves) e reutilizando conjuntos de dados existentes como o Freebase. Depois do Google, outros Grafos de Conhecimento chegaram à escala da Internet, incluindo os da Microsoft e Yahoo!

Atualmente, é a vez de empresas e administrações públicas alavancarem o conceito de Grafo de Conhecimento em nível corporativo, a fim de descrever seus dados, enriquecê-lo, interligando-o com outras bases de conhecimento dentro e fora de seu ambiente e revitalizar o desenvolvimento de sistemas intensivos em conhecimento.

Acompanhando Perez et al. (2017), comparado a 2008, o interesse nos setores de Inteligência de Mercado e de uso intensivo de dados e o papel dos Grafo de Conhecimento aumentaram drasticamente, enquanto outros, como o gerenciamento corporativo de conhecimento e o governo aberto, ainda estão lá, embora com focos ligeiramente diferentes. A seguir, são apresentados alguns aplicativos selecionados que usam Grafo de Conhecimento

em tais setores, que esperamos fornecer informações sobre o impacto potencial e as oportunidades futuras de conhecimento.

De acordo com a empresa de consultoria Internacional Data Corporation (IDC) em seu relatório de 2014 do IDG Enterprise Big Data, em média as empresas gastaram US \$ 8 milhões em alavancar o valor dos dados em 2014, com níveis de penetração de 70% e 56% para empresas e PME, respectivamente. Melhorando a qualidade do processo de tomada de decisão (59%), aumentando a velocidade dos processos de tomada de decisão (53%), melhorar o planejamento e a previsão (47%) e desenvolver novos produtos / serviços e fluxos de receita (47%) são as quatro principais áreas que aceleram o investimento em iniciativas de negócios orientadas a dados.

De acordo com Perez et al. (2017), essa tendência é especialmente acentuada no setor de conteúdo digital e publicidade. A comunicação entre marcas e consumidores está prestes a explodir. Os recursos do produto não são mais a chave para as vendas e a combinação de benefícios pessoais e coletivos está se tornando um aspecto cada vez mais crucial. De fato, as marcas que fornecem tais valores atingem um impacto maior e, conseqüentemente, obter benefícios econômicos. Por outro lado, a geração do milênio está assumindo o controle, induzindo uma mudança dramática na maneira como os consumidores e as marcas se envolvem e quais canais e tecnologias são necessários para permitir o processo.

Como resultado, os limites tradicionais da indústria da mídia estão sendo ampliados e, novas ideias, invenções e tecnologias são necessárias para acompanhar os desafios levantados pelas crescentes demandas desse mercado com uso intensivo de dados, pontual, personalizado e próspero.

Conforme Perez et al. (2017), o HAVAS, o quarto maior grupo de mídia do mundo, procura interconectar startups, inovadores, tendências tecnológicas, outras empresas e universidades em todo o mundo em uma das primeiras aplicações dos princípios de Grafo de Conhecimento em escala da Web para o mundo corporativo e a mídia. O Grafo de Conhecimento corporativo resultante suporta análises e tomada de decisões estratégicas para a incorporação de tal talento nos primeiros 18 meses de vida. Esse esforço envolve a aplicação de tecnologias semânticas, extraindo informações de inicialização de fontes on-line, estruturando-as e enriquecendo-as em um Grafo de Conhecimento autossustentável, acionável

e fornecendo às empresas de mídia conhecimento estratégico sobre as inovações mais populares.

Embora as histórias de sucesso anteriores tratem do gerenciamento do conhecimento corporativo nas empresas, neste caso, o foco está na criação de inteligência competitiva. Como já sabemos, a inovação é muitas vezes incompreendida e difícil de integrar na mentalidade corporativa e na cultura. Então, por que não ativar talentos e recursos externos relevantes quando necessário? Porém, a descoberta e a vigilância de tendências e talentos no ecossistema de startups podem ser demoradas.

O Grafo de Conhecimento, de acordo com Perez et al, (2017) da HAVAS define sua engenharia semântica para um monitoramento de vigilância da pegada digital empresarial, coletando e coletando informações e informações proveitosas, que fornece à equipe pistas claras para análise. Ao automatizar parte do processo de pesquisa, os analistas podem chegar lá com mais rapidez e precisão do que os concorrentes, aproveitando milhões de pontos de dados e implementando consistência por meio de um único ponto de entrada de conhecimento compartilhado. No momento, o Grafo de Conhecimento está sendo aberto à rede da HAVAS, com equipes em 120 escritórios em todo o mundo e clientes, fornecendo acesso ao conhecimento sobre os melhores talentos da classe para criar novas soluções de pensamento e de ponta para os desafios intermináveis e em constante evolução no setor de mídia. Com base no Grafo de Conhecimento, as equipes também avaliam e compartilham experiências, garantindo que o aprendizado possa ser propagado pela rede.

## 5. CONSIDERAÇÕES FINAIS

A comunicação entre marcas e consumidores está demandando esforços diários em função da celeridade dos atuais processos decorrentes da mudança de postura seja na forma em que se buscam informações, sejam nos perfis de usuários que tornam as buscas cada vez mais céleres. De fato, as marcas que fornecem esse valor para que se tenha um impacto mais alto e, conseqüentemente, derivará benefícios. Por outro lado, a geração do milênio está assumindo o controle; induzindo uma mudança dramática na maneira como consumidores e marcas se envolvem e quais canais e tecnologias são necessários para ativar o processo. Como resultado, os limites tradicionais da mídia da indústria está sendo ampliada e novas ideias, invenções e tecnologias são necessárias para acompanhar os desafios levantados pelas demandas crescentes desse uso intensivo de dados.

Para que haja a confiabilidade e a precisão do Grafo de Conhecimento, o uso de fontes autorizadas e especializadas e a decorrente priorização da atualização sobre os volumes informacionais é necessário. Contudo, há ainda pontos de atenção a exemplo da resolução e desambiguação das entidades, especialmente quando dados não estruturados de domínios ilimitados entram em jogo. Durante a integração e enriquecimento de dados, várias entidades candidatas podem ser identificadas.

Outros desafios incluem gerenciamento de tempo e versão (sub) grafo, controle de atualizações automáticas versus humanas e resiliência contra alterações nos dados fontes, especialmente APIs da Web. É particularmente importante monitorar possíveis decaimentos em um Grafo de Conhecimento, aplicando as técnicas existentes de curadoria e métodos extraídos de domínios científicos.

O domínio aparelhado do conhecimento é uma das principais estratégias que permitem às empresas explorar plenamente o seu conhecimento coletivo. No entanto, algumas barreiras geralmente limitam o potencial desta abordagem:

- Participação forte e ativa (provisão de conhecimento);
- A falta de tecnologias inteligentes verdadeiramente evoluídas que permitam aos funcionários beneficiar do conhecimento global fornecido pelas empresas e outros usuários (consumindo conhecimento).

Empresas de consultoria internacional apontam investimentos significativos em alavancagem no valor dos dados com níveis de penetração significativos para corporações com intuito de melhorar a qualidade da tomada de decisão, aumentando a velocidade dos processos, melhorando seu planejamento, previsão, desenvolvimento de novos produtos / serviços e receita, acelerando iniciativas no investimento em negócios orientados a dados.

A abordagem clássica da Gestão do Conhecimento com soluções que são poderosas e complexas vem aos poucos seguindo a ideia de simplicidade de funcionalidades semânticas construídas sobre o Grafo de Conhecimento que objetivam a captura e relacionamento das entidades corporativas relevantes.

Os Grafos de Conhecimento fornecem novos paradigmas e plataformas para implementar uma variedade de bases de conhecimento corporativo de uma forma mais eficaz e poderosa. Para organizações empresariais, utilizar Grafos de Conhecimento significa implementar arquitetura e suporte de seus processos de aquisição e manutenção. Os limites organizacionais tornam possível (ainda que não trivial em muitos casos) o pré-requisito de suporte chave de tais processos.

Essa tendência é especialmente aguda no setor de conteúdo digital e de publicidade, onde os recursos do produto não são mais a chave para as vendas, mas sim a combinação de benefícios pessoais e coletivos que está se tornando um aspecto cada vez mais crucial. O Grafo de Conhecimento corporativo agrega em seu compromisso o suporte de análises e tomada de decisões estratégicas que envolvem a aplicação de tecnologias, extraindo informações, estruturando e enriquecendo um portfólio de conhecimento acionável e autossustentável, mesclando as histórias de sucesso que lidam com a gestão do conhecimento corporativo dentro das corporações, com o foco na criação de inteligência competitiva de forma colaborativa e disruptiva.

Nesta pesquisa, pretendeu-se avaliar o estado de utilização dos Grafos de Conhecimento: incitar insights acerca dos principais desafios na construção, gerenciamento e armazenamento dos Grafos de Conhecimento bem como, possíveis abordagens para tratar questões controversas que envolvem sua utilização.

A construção de um Grafo de Conhecimento não é o tópico central desta pesquisa, porém, dada a matéria prima envolvida, ou seja, a Informação é imprescindível que para a Ciência da Informação que sua multidisciplinariedade encontre contornos possíveis de

atuação no contexto, dada também a gama de ferramentas ora abordadas as quais são parte de seu *expertise*.

Notadamente, temos também a consideração quanto às transições de cunho social, notadamente a exponencial necessidade de gestão do universo informacional ora criado, que carece de tratamento e visão emergente que acompanhe tendências tecnológicas como o Big Data, por exemplo, que, nas palavras de Davenport (2017), big data é um termo genérico para dados que não podem ser contidos nos repositórios usuais; refere-se a dados volumosos demais para caber em um único servidor; não estruturados demais para se adequar a um banco de dados organizados em linhas e colunas; ou fluídos demais para serem armazenados em um *data warehouse* estático. Embora o termo enfatize seu tamanho, o aspecto mais complicado do big data, na verdade, envolve sua falta de estrutura. Sendo assim, conclui-se que é necessário desconstruir e desmistificar certos termos através do arcabouço conceitual que é detido pela Ciência da Informação e lançar luz sobre possíveis transformadores de enorme potencial, seja para sociedade, seja para o mundo corporativo.

Nos últimos anos, os Grafos de Conhecimento surgiram como um dos sistemas mais importantes de representação do conhecimento, organização e compreensão. Um vasto número de pesquisas questões, aplicações e produtos relacionados aos Grafos de Conhecimento foram propostos, mas é consenso que ainda existem muitos desafios e oportunidades no campo dos Grafos de Conhecimento.

Inspirados na história de sucesso do Google, os Grafos de Conhecimento estão ganhando força na a World Wide Web. Nos últimos anos, o esforço contínuo realizado em pesquisa industrial decorreu em uma crescente aceitação industrial por outros gigantes da Internet, que incluem a pesquisa de grafos do Facebook e Satori da Microsoft.

Recursos adicionais, novas ideias e várias aplicações foram introduzidos e, como consequência, a noção de grafos de conhecimento cresceu de forma ampla, englobando toda uma linha de esforço da comunidade por conta própria desenvolver novos métodos e tecnologias.

No entanto, há uma nova tendência que ganha impulso, que se encontra na percepção de que esse valor é diretamente proporcional à interligação dos dados não apenas em sistemas complexos e abertos, como a Web, mas também em empresas específicas, aplicativos baseados em combinações de dados corporativos e abertos. Mais adequado para operações de

busca e relacionamento, dados mal formalizados, mas altamente interconectados estão se tornando mais populares do que conjuntos de dados altamente formalizados, mas isolados. O atual cenário de aplicativos, mais orientado para dispositivos móveis e em tempo real, está impondo essa nova mudança de paradigma.

As técnicas envolvidas nos Grafos de Conhecimento são compostas de novas abordagens para lidar com os desafios interação em volumes informacionais enfrentados por organizações. Fornecer detalhes técnicos necessários para os profissionais em potencial nessas organizações é tarefa da Ciência da Informação aliada a outras áreas de conhecimento as quais conjugam aportes teóricos necessários para compreendê-los e consumi-los em suas formas mais diversas. Desse modo, é necessária para soluções que envolvam os Grafos de Conhecimento, profundidade teórica que dará suporte a modelos de dados RDF, Linguagens de ontologia OWL e dados vinculados. Uma visão clara embasada nos sistemas de organização do conhecimento é também essencial para a construção de um Grafo de Conhecimento bem-sucedido e a manutenção de um ecossistema saudável e sustentável.

Os Grafos de Conhecimentos se apresentam como ferramentas poderosas, porém, ainda precisamos superar o obstáculo de qualidade de dados, ou seja, avaliar e validar a confiabilidade de diferentes fontes de informação, de modo a garantir uma compreensão valiosa de fenômenos complexos e mecanismos que sejam derivados de preceitos da organização da informação de modo a direcioná-los no sentido da inovação tornando-os elementos confiáveis de aferição da informação. Na verdade, estamos apenas começando a reunir experiência em como integrar verdadeiramente fontes tão diferentes, em como manter conjunto de dados tão volumoso, como garantir a qualidade, correção e cobertura de tais grafos, como processá-los de forma eficiente e conseguir consolidar uma governança eficaz.

Os Grafos de Conhecimento estão sendo disseminados de forma exponencial, determinando uma situação complexa e provocando a área de organização do conhecimento no sentido de prover respostas a questões de ordem emergencial, pois, além de problemas técnico científicos, há um problema político-social, devido ao fato de que os maiores Grafos de Conhecimento são controlados por empresas, que normalmente não atribuem a proveniência do conhecimento em qualquer fragmento de um Grafo de Conhecimento determinando pontos de atenção. Isso deve ser uma preocupação para todos os desenvolvedores usuários e padrões de proveniência e confiança apropriados devem ser aplicados ou recomendados, sejam eles formalmente expressos ou não. Muitos desafios

surgirão na construção do conhecimento baseado em grafos, porém, dado a variedade das tendências tecnológicas emergentes, sugere-se que mecanismos de indexação e recuperação estejam sensíveis a responder de forma confiável às requisições de informações.

Vale ressaltar que o Google Knowledge Graph, é uma base de conhecimento do sistema de pesquisa da empresa Google que visou melhorar os resultados da sua ferramenta de busca com informações de pesquisa semântica. Internamente possui o Googlebot que é o robô indexador usado pelo Google para indexar páginas e documentos da web para construir um índice de buscas para o Google Search tornando-o mais eficiente.

## 6. REFERENCIAS

- STEFANINI, G. Afinal, o que é a Inovação Aberta e quais são seus benefícios? Disponível em <<https://stefanini.com/pt-br/trends/artigos/beneficios-da-inovacao-aberta>> Acesso em :01 Ago. 2019.
- ALMEIDA, Maurício Barcellos (2014). Uma abordagem integrada sobre ontologias: Ciência da Informação, Ciência da Computação e Filosofia. Disponível em <<http://www.scielo.br/pdf/pci/v19n3/a13v19n3.pdf>> Acesso em :05 Jul. 2019.
- ALVARENGA, L. Representação do conhecimento na perspectiva da Ciência da Informação em tempo e espaço digitais. 2003. Disponível em: <<http://www.periodicos.ufsc.br/index.php/eb/article/view/97>>. Acesso em: 21 mar. 2019.
- ALVARES, Lilian (org.) Organização da informação e do conhecimento: conceitos, subsídios interdisciplinares e aplicações. São Paulo: B4Editores, 2012.
- BLAIR, David C. Information Retrieval and the Philosophy of Language. En: Blaise Cronin (Ed.): Annual Review of Information Science and Technology. 2003. Disponível em <[https://is.muni.cz/el/1421/jaro2011/VIKBA12/um/Bates\\_2005\\_intro.pdf](https://is.muni.cz/el/1421/jaro2011/VIKBA12/um/Bates_2005_intro.pdf)> Acesso em:24 set. 2019.
- BORKO, H. Information Science: What is it? American Documentation, v.19, n.1, p.3-5, Jan. 1968. (Tradução Livre). Disponível em: <[https://edisciplinas.usp.br/pluginfile.php/3433774/mod\\_resource/content/1/Oque%C3%A9C\\_L.pdf](https://edisciplinas.usp.br/pluginfile.php/3433774/mod_resource/content/1/Oque%C3%A9C_L.pdf)>. Acesso em 21 mar. 2019.
- CAFÉ, L.; BRASCHER, M. Organização da Informação ou Organização do Conhecimento?. IX ENANCIB 28 de setembro de 2008 – 1 de outubro de 2008. São Paulo. 2008. Disponível em <<http://enancib.ibict.br/index.php/enancib/ixenancib/paper/viewFile/3016/2142>> Acesso em 01-jun-2020.
- CAPURRO, R.; HJØRLAND, B. The concept of information. Annual Review of Information Science of Technology, New York, v. 37, p. 343-411, 2003. Disponível em: <<http://www.capurro.de/infoconcept.html>>. Acesso em: 21 set. 2019.
- CASTELLS, Manuel. A Sociedade em Rede. 11 ed. São Paulo: Paz e Terra, 1999
- CERVO, A. L.; BERVIAN, P. A. Metodologia científica. 5. ed. São Paulo: Prentice Hall, 2003.
- CHESBROUGH, H. Inovação aberta: como criar e lucrar com a tecnologia. Porto Alegre: Bookman, 2012.
- COSTA, M. A. F.; COSTA, M. de F. B. Metodologia da pesquisa: conceitos e técnicas. Rio de Janeiro: Interciência, 2001.
- DAHLBERG, I. Knowledge organization: a new science? Knowledge Organization, Frankfurt, v.33, n.1, p. 11-19, 2006.

EHRLINGER, L. and Wolfram Wöß, Towards a Definition of Knowledge Graphs, Institute for Application Oriented Knowledge Processing. Disponível em <[https://www.researchgate.net/profile/Wolfram\\_Woess/publication/323316736\\_Towards\\_a\\_Definition\\_of\\_Knowledge\\_Graphs/links/5a8d6e8f0f7e9b27c5b4b1c3/Towards-a-Definition-of-Knowledge-Graphs.pdf](https://www.researchgate.net/profile/Wolfram_Woess/publication/323316736_Towards_a_Definition_of_Knowledge_Graphs/links/5a8d6e8f0f7e9b27c5b4b1c3/Towards-a-Definition-of-Knowledge-Graphs.pdf)>. Acesso em: 12 abr. 2019.

ALDECOA , Marcel Castellani. Enterprise 2.0: O impacto da Web 2.0 nas organizações e nos tradicionais modelos de gestão. Tiespecialistas. São Paulo, 7 de mai 2012. Disponível em <<https://www.tiespecialistas.com.br/enterprise-2-0-o-impacto-da-web-2-0-nas-organizacoes-e-nos-tradicionais-modelos-de-gestao>> Acesso em 01 Ago. 2019.

ESTEBAN NAVARRO, M.A. El marco disciplinar de los lenguajes documentales: la Organización del Conocimiento y las ciencias sociales. Scire, Zaragoza, v. 2, n. 1, jun., 1996.

FURNIVAL, A. C. M. ; COSTA, L. S. F. (Org.) . Informação e conhecimento - aproximando áreas de saber. 1. ed. São Carlos: EDUFSCar, 2005.

GIL, A. C. Como elaborar projetos de pesquisa. 4. ed. São Paulo: Atlas, 2007.

International Society for Knowledge Organization. Disponível em <[https://www.isko.org/cyclo/knowledge\\_organization#ref](https://www.isko.org/cyclo/knowledge_organization#ref) > Acesso em: 28 Jan 2020.

GOMES, H. E. . Tendências da pesquisa em organização do conhecimento. 2011. (Apresentação de Trabalho/Comunicação). Disponível em <<http://www.brapci.inf.br/index.php/article/download/6974>> Acesso em 5 Fev 2020.

J. C. Binwal - Ranganathan and the Universe of Knowledge Disponível em <<https://www.nomos-elibrary.de/10.5771/0943-7444-1992-4-195/ranganathan-and-the-universe-of-knowledge-jahrgang-19-1992-heft-4>> Acesso em 28 Jan 2020.

KOBASHI, Nair Yumiko. Fundamentos semânticos e pragmáticos da construção de instrumentos de representação de informação. Pesquisa Brasileira em Ciência da Informação e Biblioteconomia, v. 3, n. 1, 2008.

KÖCHE, J. C. Fundamentos de metodologia científica: teoria da ciência e iniciação à pesquisa. Petrópolis, RJ: Vozes, 1997.

LE COADIC, Y. F. A Ciência da Informação. 2. ed. Brasília: Briquet de Lemos, 2004.

LOUREIRO et al. (2014). Indução de Árvores de Decisão na Descoberta de Conhecimento: Caso de Empresa de Organização de Eventos. 10.13140/2.1.2190.1121. Disponível em <[https://www.researchgate.net/publication/270449012\\_Inducao\\_de\\_Arvores\\_de\\_Decisao\\_na\\_Descoberta\\_de\\_Conhecimento\\_Caso\\_de\\_Empresa\\_de\\_Organizacao\\_de\\_Eventos](https://www.researchgate.net/publication/270449012_Inducao_de_Arvores_de_Decisao_na_Descoberta_de_Conhecimento_Caso_de_Empresa_de_Organizacao_de_Eventos)> Acesso em: 22 abr. 2019.

MCAFEE, Andrew. Empresas 2.0: a força das mídias colaborativas para superar grandes desafios empresariais. Rio de Janeiro: Elsevier, 2010.

OSTROSKI, Alvaro. Teoria dos Grafos e Aplicações. 2009. Disponível em <<http://revistas.utfpr.edu.br/pb/index.php/SysScy/article/view/709>> Acesso em: 22 abr. 2019.

PAULHEIM, Heiko. (2016). Knowledge graph refinement: A survey of approaches and evaluation methods. *Semantic Web*. 8. 489-508. 10.3233/SW-160218. Disponível em < [https://www.researchgate.net/publication/311479070\\_Knowledge\\_graph\\_refinement\\_A\\_survey\\_of\\_approaches\\_and\\_evaluation\\_methods/citation/download](https://www.researchgate.net/publication/311479070_Knowledge_graph_refinement_A_survey_of_approaches_and_evaluation_methods/citation/download) > Acesso em: 25 Ago. 2019.

PEREZ et al. *Exploiting Linked Data and Knowledge Graphs in Large Organizations*, Switzerland: Springer International Publishing, 2017.

RAMALHO, R.A.S. Desenvolvimento e utilização de ontologias em Bibliotecas Digitais: uma proposta de aplicação, p. 18, 2010. Disponível em < [https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho\\_ras\\_do\\_mar.pdf](https://www.marilia.unesp.br/Home/Pos-Graduacao/CienciadaInformacao/Dissertacoes/ramalho_ras_do_mar.pdf) > Acesso em : 9 set. 2017.

RAMALHO, R.A.S. Análise do Modelo de Dados SKOS: Sistema de Organização do Conhecimento Simples para a Web. *Informação & Tecnologia (Itec)*, v. 2, p. 68, 2015. Disponível em: <[http://periodicos.ufpb.br/ojs/index.php/itec/art2016\\_icle/view/25995](http://periodicos.ufpb.br/ojs/index.php/itec/art2016_icle/view/25995)> Acesso em: 9 set. 2017.

RIBEIRO, Fernanda. Organizar e representar informação: apenas um meio para viabilizar o acesso? Porto, 25 mai 2005. Disponível em <<http://ler.letras.up.pt/uploads/ficheiros/1239.pdf>> Acesso em: 26 Jan 2020.

RIGNEL et. al. UMA INTRODUÇÃO A LÓGICA FUZZY. *Revista Eletrônica de Sistemas de Informação e Gestão Tecnológica*, v. 1, n.1, mar. 2011. Disponível em <[http://www.logicafuzzy.com.br/wp-content/uploads/2013/04/uma\\_introducao\\_a\\_logica\\_fuzzy.pdf](http://www.logicafuzzy.com.br/wp-content/uploads/2013/04/uma_introducao_a_logica_fuzzy.pdf)>. Acesso em: 25 Jun. 2019.

RODRIGUES, Alexander. Otimização Semântica para buscas. *Semântico*. São Paulo, 2 de jun de 2019. Disponível em <https://semantico.com.br/otimizacao-semantica-para-buscas/>. Acesso em 01-jun-2020.

RODRIGUES, Aleander .O poder do Grafo do Conhecimento. *Semântico* ,São Paulo, 14 de ago de 2019. Disponível em <<https://semantico.com.br/o-poder-do-grafo-do-conhecimento>> Acesso em 01-jun-2020.

ROGERS, David L., *Transformação digital: repensando seu negócio para a era digital*. Ed.1. São Paulo: Autentica Business, 2018.

SAUSSURE, Ferdinand de. *Curso de Linguística Geral*. São Paulo: Cultrix, 1975 disponível em < <https://joacamillopenna.files.wordpress.com/2016/04/saussurre-curso-de-linguistica-geral.pdf> > Acesso em 26 Set. 2019.

SILVEIRA, D. T.; CÓRDOVA, F. P. A pesquisa científica. In: GERHARDT, T. E.; SILVEIRA, D. T. (Orgs.). *Métodos de pesquisa*. Porto Alegre: Editora da UFRGS, 2009.

SANTOS, Plácida Leopoldina Ventura Amorim da Costa. Catalogação, formas de representação e construções mentais. *Tendências da Pesquisa Brasileira em Ciência da Informação*, v. 6, n. 1, p. 1-24, 2013. Disponível em: <<http://hdl.handle.net/11449/115044>>. Acesso em: 22 Jan 2020.

SOUZA, J. L.; MARTINS, P. G. M.; RAMALHO, R. A. S. Modelos de representação semântica na era do big data. *Brazilian Journal of Information Science*, v. 12 No 3, n. 3, p. 34 ao 40, 2018.

SOUZA, Renato Rocha; ALMEIDA, Maurício Barcellos; BARACHO, Renata Maria Abrantes. *Ciência da Informação em transformação: Big Data, Nuvens, Redes Sociais e Web Semântica*. *Ciência da Informação*, [S.l.], v. 42, n. 2, aug. 2015. ISSN 1518-8353. Disponível em: <<http://revista.ibict.br/ciinf/article/view/1379>>. Acesso em: 25 mar. 2019.

SOUZA, Rosali Fernandez de. Organização e representação de áreas do conhecimento em ciência e tecnologia: princípios de agregação em grandes áreas segundo diferentes contextos de produção e uso de informação. *Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação*, Florianópolis, p. 27-41, jul. 2006. ISSN 1518-2924. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2006v11nesp1p27/384>>. Acesso em: 05 fev. 2020.

SOWA, John. (2000). *Knowledge Representation: Logical, Philosophical, and Computational Foundations*. Disponível em <<https://www.researchgate.net/publication/225070439>> Acesso em : 22 Ago. 2019.

TERRA, J.C.C. Gordon, C. *Portais corporativos: A revolução na gestão do conhecimento*. São Paulo: Negócio Editora, 2002.

YAN, Jihong ; Wang, Chengyu & Cheng, Wenliang & Gao, Ming & Zhou, Aoying. 2016). A retrospective of knowledge graphs. *Frontiers of Computer Science*. 2016. Disponível em <[https://www.researchgate.net/publication/308691086\\_A\\_retrospective\\_of\\_knowledge\\_graph\\_s](https://www.researchgate.net/publication/308691086_A_retrospective_of_knowledge_graph_s)>. Acesso em: 25 Ago. 2019.