

XX ENANCIB

21 a 25 Outubro/2019 – Florianópolis

A Ciência da Informação e a era da Ciência de Dados

ISSN 2177-3688

GT-8 - Informação e Tecnologia

PAPÉIS DO PROFISSIONAL DA INFORMAÇÃO NA EVOLUÇÃO DA WEB SEMÂNTICA

ROLES OF THE INFORMATION PROFESSIONAL IN THE SEMANTIC WEB EVOLUTION

Clayton Martins Pereira - Universidade Estadual Paulista Júlio de Mesquita Filho

Larissa Pavarini da Luz - Universidade Estadual Paulista Júlio de Mesquita Filho

José Eduardo Santarem Segundo - Universidade Estadual Paulista Júlio de Mesquita Filho

Modalidade: Trabalho Completo

Resumo: Atualmente nos encontramos numa fase de transição de uma Web de documentos, onde os conteúdos publicados podem ser recuperados e aproveitados somente por humanos, para uma Web de dados, que utiliza a Web Semântica, onde os conteúdos publicados podem ser recuperados e aproveitados tanto por humanos quanto por máquinas, provocando uma revolução na forma de gerar e consumir informação nos dias atuais. O objetivo deste trabalho é apresentar um panorama quantitativo das pesquisas já realizadas neste tema nos últimos dez anos, com ênfase nas pesquisas realizadas dentro da Ciência da Informação, além de um rol de atividades relacionadas à publicação de dados abertos ligados e ao desenvolvimento de aplicações para consumo destes dados, nas quais o profissional da informação pode se sobressair e contribuir de forma significativa para a evolução da Web Semântica. A partir de um levantamento bibliométrico nas bases de dados *Web of Science* e *Scopus* foi possível apontar que as pesquisas relacionadas à Web Semântica estão em crescente evolução nos últimos dez anos, porém ainda fortemente concentradas na Ciência da Computação. Entretanto, as pesquisas sobre este tema na área da Ciência da Informação, apesar de ainda pouco expressivas em termos de quantidade, apresentam uma tendência de crescimento nos últimos cinco anos. Por fim, espera-se que este trabalho sirva de incentivo para que os atuais e futuros profissionais da informação possam contribuir com a evolução e consolidação da Web Semântica e, conseqüentemente, da Web de dados, revolucionando a forma como consumimos informação nesta era da ciência de dados.

Palavras-Chave: Ciência da Informação; Profissional da Informação; Mapeamento de Competências; Web de Dados; Web Semântica.

Abstract: We are in the process of transitioning from a Web of documents, where the published contents can be recovered and used only by humans, to a Web of data, which makes use of the Semantic Web, where the published contents can be recovered and used not only by humans but also by machines, revolutionizing the way one can generate and consume information nowadays. The objective of this work is to present a quantitative outlook on the researches addressing this theme over the past ten years, emphasizing the researches carried out in the Information Science field, as well as a list of activities related to the publishing of linked open data and the development of applications for the consumption of these data, in which the professional of information can stand out and contribute significantly for the evolution of the Semantic Web. From a bibliometric survey on the databases *Web of Science* and *Scopus*, it was possible to point out that the researches related to the

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

Semantic Web have been constantly evolving over the past ten years, but they are substantially concentrated in Computer Science. Although the researches on this theme in the field of Information Science are not expressive in numbers, they have been growing during the last five years. This work is expected to become an incentive for the current and future professionals of information contributing with the evolution and consolidation of the Semantic Web and, consequently, the Web of Data, revolutionizing the way we consume information in this era of data science.

Keywords: Information Science; Professional of Information; Competences Mapping; Web of Data; Semantic Web.

1 INTRODUÇÃO

No presente momento nos encontramos em plena era da ciência de dados, que segundo Amaral (2016) “é uma ciência que trata de dados para obter conhecimento e informação de forma sistemática, que normaliza e organiza esse conhecimento.”

Trata-se de uma era em que tal organização do conhecimento pede uma mudança, que é o fato da transição de uma Web baseada em documentos, na qual os conteúdos publicados podem ser recuperados e aproveitados somente por humanos, para uma Web baseada em dados, na qual os conteúdos publicados podem ser recuperados e aproveitados tanto por humanos quanto por máquinas, revolucionando assim a forma de gerar e consumir informação.

Para que tal transição seja bem sucedida e a Web de dados seja uma realidade presente em nosso cotidiano, se fazem necessários a pesquisa e o desenvolvimento contínuo de uma Web Semântica onde as informações possam ser disponibilizadas em formato aberto (bruto) e a combinação dessas informações gerem significados bem definidos, ou seja, exatamente de acordo com o contexto ou domínio de conhecimento ao qual estejam aplicadas (semântica), de forma a possibilitar que além de humanos, máquinas sejam capazes de processá-las e entendê-las automaticamente (BERNERS-LEE; HENDLER; LASSILA, 2001). Isso permitirá que tais informações sejam combinadas e compartilhadas a partir de diferentes fontes, além de habilitar seu uso em diferentes contextos por diversas aplicações e agentes de software (LV; MA, 2008). Desta forma, o sucesso da Web Semântica depende da criação massiva de dados abertos e do conhecimento que isso venha agregar no uso ou aplicabilidade da informação gerada (LING; ZHOU, 2010).

No ano de 2009, Tim Berners-Lee (considerado o pai da Web), em palestra concedida para a TED¹, já pedia o compartilhamento massivo de dados brutos por todas as pessoas ao redor do mundo. Berners-Lee, naquela época, planejava o que hoje encontra-se em fase de uso: uma Web para dados abertos e interligados por meio de uma Web Semântica, uma rede que pode fazer pelos dados aquilo que a World Wide Web fazia pelos documentos, imagens e vídeos, ou seja, permitir ampla disponibilidade e novas maneiras de avançar o conhecimento.

Nesse contexto, o objetivo deste trabalho é apresentar um panorama quantitativo das

¹ TED é uma organização não-partidária, sem fins lucrativos, dedicada a disseminar ideias por meio de conferências, que hoje abrange os mais variados temas: ciência, tecnologia, negócios, questões globais, entre outros.

pesquisas já realizadas sobre a Web Semântica nos últimos dez anos (2009 a 2018), enfatizando as pesquisas realizadas no âmbito da Ciência da Informação. A partir deste panorama será possível prospectar temáticas de pesquisa nas quais o profissional da informação pode se sobressair, dadas as competências adquiridas durante sua formação, e que podem contribuir significativamente para a evolução da Web Semântica e do profissional da informação.

A metodologia empregada neste trabalho, de natureza quali-quantitativa e de tipo descritivo, consiste primeiramente na realização de uma pesquisa bibliográfica sobre o tema Web Semântica, de forma a permitir uma breve revisão sobre o assunto, seguido de um levantamento bibliométrico dos trabalhos sobre este tema, publicados em periódicos indexados nas bases de dados *Web of Science* e *Scopus*, no período de 2009 a 2018, o qual concentra a maior quantidade de trabalhos publicados sobre o tema. Por fim, a partir de uma análise documental, foram identificados quais assuntos sobre a Web Semântica ainda têm sido pouco explorados na área da Ciência da Informação, e que podem significar oportunidades para a atuação do profissional da informação.

O presente trabalho está estruturado da seguinte forma: a seção 2 apresenta uma revisão bibliográfica do tema Web Semântica e como se comporta o contexto da atual Web de Dados. Na seção 3 são apresentados os levantamentos bibliométricos relacionados aos trabalhos sobre o tema, publicados em periódicos indexados nas bases de dados *Web of Science* e *Scopus*, com destaque para aqueles publicados em periódicos da área de Ciência da Informação. Por fim, são apresentadas as perspectivas para os profissionais da informação nas pesquisas sobre o tema, e sua importância na contextualização e criação da Web de Dados.

2 WEB SEMÂNTICA: DA WEB DE DOCUMENTOS PARA A WEB DE DADOS

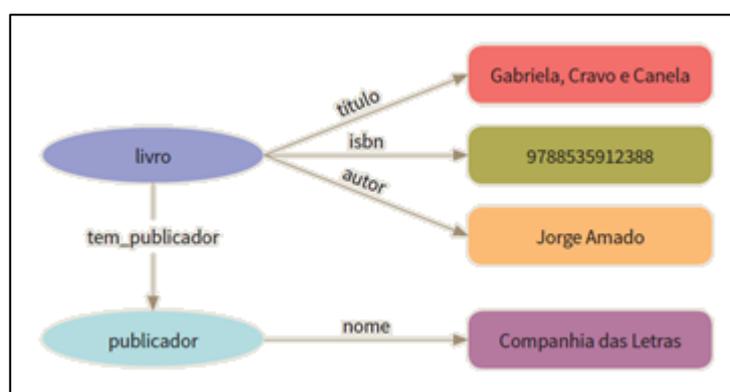
A Web atual, de uma forma geral, tem seu conteúdo destinado à leitura humana (foco no nível de apresentação dos dados), o que impossibilita a manipulação por computadores. Mas então por que precisamos tanto da Web Semântica atualmente? Ou então por que a Web de hoje não é suficiente?

Marissa Mayer, vice-presidente de produtos de pesquisa e experiência do usuário da Google, em uma entrevista para o jornal *The Guardian*, disse enfaticamente que “dada a escala da Web e o poder do computador que temos hoje, estamos começando a ver coisas

(recursos) e as relações entre eles, que podem ser representadas em uma sintaxe XML (*eXtensible Markup Language*, ou Linguagem de Marcação Estendida), a qual permite a troca de dados e metadados entre aplicações. Este modelo permite a codificação, a troca e o reuso de dados e informações, por meio de uma representação estruturada, onde significados são codificados, de maneira uniforme, em conjuntos de triplas objeto-propriedade-valor chamadas de afirmações, tal como numa sentença composta por sujeito-predicado-objeto, de forma a facilitar sua manipulação por máquinas (ANTONIOU E HARMELEN, 2008).

A Figura 2 mostra um exemplo de um grafo RDF que relaciona um recurso [livro] (com três triplas de propriedade e objeto para este sujeito) a um recurso [publicador] (com uma tripla de propriedade e objeto para este sujeito).

Figura 2: Exemplos de triplas RDF e de relacionamento entre recursos.



Fonte: Laufer (2015).

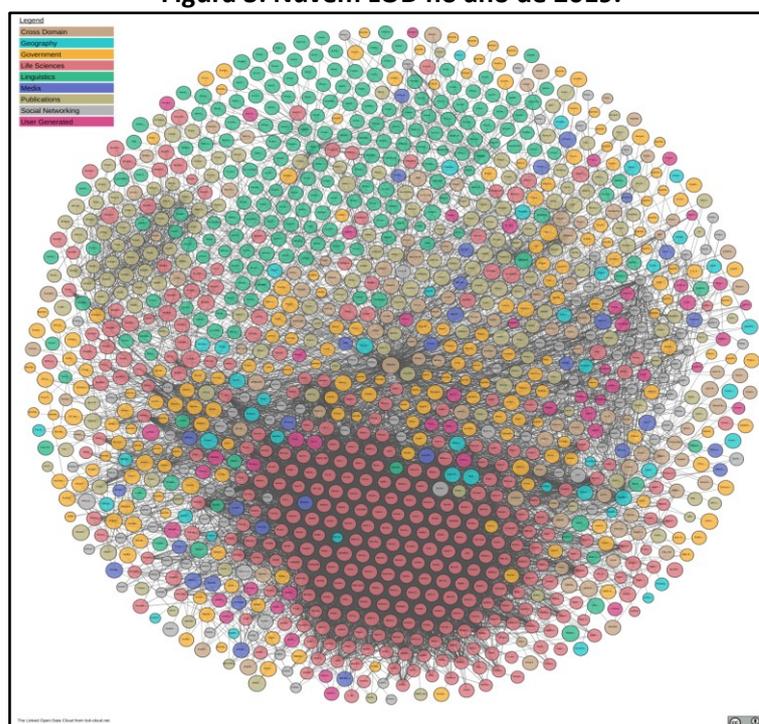
As ontologias consistem em estruturas formais de conceitos e relações entre conceitos, que tipicamente incluem hierarquias de classes, além de um conjunto de axiomas que restringe a interpretação dessa estrutura, com a finalidade de definir uma especificação formal de um domínio de conhecimento, por meio de descrições de classes, propriedades e relações entre estas (ALLEMAND E HENDLER, 2008). Assim, as ontologias assumem um importante papel no compartilhamento do conhecimento e na interoperabilidade entre sistemas, ao prover uma forma de tornar os dados legíveis e entendíveis por máquina (MOHAMED; JINCAI; QIAN, 2010).

Os usuários podem gravar conceituações explícitas e formais de modelos de domínio, por meio de linguagens ontológicas, tais como a RDFS (*Resource Description Framework Schema*, ou Estrutura do Modelo de Descrição de Recursos), uma linguagem universal que

permite aos usuários descrever recursos utilizando seus próprios vocabulários (ANTONIOU E HARMELEN, 2008), e a OWL (*Ontology Web Language*, ou Linguagem Ontológica da Web), uma linguagem de marcação semântica para publicar e compartilhar ontologias na Web por meio da descrição de classes, propriedades, e das relações entre estas, além de possibilitar a definição de características como restrições, afirmações disjuntas e especificações de relacionamentos lógicos entre objetos (ALLEMAND E HENDLER, 2008).

Já os dados ligados (*linked data*) são uma das mais bem-sucedidas tecnologias da Web Semântica. Baseado nos seus princípios, surgiu o projeto Linked Open Data (LOD), cujo principal objetivo é criar um espaço global de dados que sejam processáveis por máquinas, onde um crescente número de *datasets* (conjuntos de dados) possam ser disponibilizados e ligados entre si, sem intervenção humana, de forma a tornar realidade a Web de dados (HERT *et al.*, 2012). A Figura 3 ilustra uma nuvem de conexões (*links*) entre os conjuntos de dados abertos (nós) disponibilizados neste espaço, a chamada nuvem LOD.

Figura 3: Nuvem LOD no ano de 2019.



Fonte: lod-cloud.net (2019).

Os dados ligados podem ser gerados a partir da disponibilização de *datasets*, na forma de triplas RDF, acessíveis por meio de Identificadores Únicos de Recursos (URI, similar a uma URL, *Uniform Resource Locator*, que é o endereço de um recurso disponível em uma rede,

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

como por exemplo o localizador padrão de recursos da Internet HTTP, *Hypertext Transfer Protocol*, protocolo de comunicação utilizado para sistemas de informação de hipermídia, distribuídos e colaborativos), e que podem ser recuperados a partir de uma linguagem de consulta para RDF (SPARQL, considerada a linguagem SQL da Web Semântica) (MACHADO E PARENTE, 2011).

Assim, a linguagem SPARQL permite a consulta de triplas RDF obtidas a partir de *datasets* ou de *triple stores* (depósitos de triplas). Para isso, é baseada em um modelo de busca onde a expressão de consulta é montada na forma de triplas, mas com a possibilidade de definir variáveis em quaisquer dos termos desta tripla, seja no sujeito, no predicado ou no objeto. A Figura 4 mostra um exemplo de consulta SPARQL, e seu respectivo resultado, efetuada sobre um conjunto de triplas.

Figura 4: Exemplo de consulta SPARQL.

Conjunto de triplas	Consulta SPARQL						
<pre>base <http://example.org/> . @prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> . @prefix foaf: <http://xmlns.com/foaf/0.1/> . @prefix dc: <http://purl.org/dc/elements/1.1/> . <#gabriela-cravo-canela> dc:identifier "9788535912388" ; foaf:name "Gabriela, Cravo e Canela" ; dc:creator "Jorge Amado" ; dc:publisher <#companhia-das-letras> . <#vidas-secas> dc:identifier "9788501067340" ; foaf:name "Vidas Secas" ; dc:creator "Graciliano Ramos" ; dc:publisher <#grupo-record> . <#antologia-poetica> dc:identifier "9788535921199" ; foaf:name "Antologia Poética" ; dc:creator "Carlos Drummond de Andrade" ; dc:publisher <#companhia-das-letras> . <#companhia-das-letras> foaf:name "Companhia das Letras". <#grupo-record> foaf:name "Grupo Editorial Record" .</pre>	<pre>PREFIX dc: <http://purl.org/dc/elements/1.1/> PREFIX foaf: <http://xmlns.com/foaf/0.1/> PREFIX ex: <http://example.org/> SELECT ?name ?creator WHERE { ?book dc:publisher ex:companhia-das-letras . ?book foaf:name ?name . ?book dc:creator ?creator . }</pre>						
Resultado retornado							
<table border="1"> <thead> <tr> <th align="center">NAME</th> <th align="center">creator</th> </tr> </thead> <tbody> <tr> <td align="center">"Gabriela, Cravo e Canela"</td> <td align="center">"Jorge Amado"</td> </tr> <tr> <td align="center">"Antologia Poética"</td> <td align="center">"Carlos Drummond de Andrade"</td> </tr> </tbody> </table>		NAME	creator	"Gabriela, Cravo e Canela"	"Jorge Amado"	"Antologia Poética"	"Carlos Drummond de Andrade"
NAME	creator						
"Gabriela, Cravo e Canela"	"Jorge Amado"						
"Antologia Poética"	"Carlos Drummond de Andrade"						

Fonte: Autores, adaptado de Laufer (2015).

A linguagem SPARQL abre caminho para que uma expressão de consulta venha a ser redigida de uma forma muito próxima à linguagem natural. Os resultados obtidos nas

consultas podem ainda assumir a forma de dados ligados, os quais podem ser explorados levando a outros resultados não obtidos inicialmente.

A partir desse contexto computacional, verifica-se uma evolução de como todo esse processo impacta em nossas vidas, como está sendo sua ascensão e ainda como ela caminha para que venha existir de forma total, ou seja, fazer com que toda essas tecnologias venham a trazer melhores resultados, facilitando o máximo da execução de tarefas, quer seja para humanos como para máquinas.

Por isso, a Web Semântica, segundo Berners-Lee, Hendler e Lassila (2001), é apresentada como uma extensão da Web atual, pois tem muito mais o propósito de gerar conhecimento, ou seja, está ligado muito mais na relação e na reutilização, do que somente a produção em massa de conteúdo.

Além do que a premissa para que tudo isso aconteça é a necessidade de uma estrutura padronizada, como já apresentado, para que além de humanos, as máquinas, sistemas, robôs, entre outros, consigam atingir essa essência, da real utilização dos dados, podendo assim interpretar as informações.

3 PANORAMA GERAL DAS PESQUISAS SOBRE WEB SEMÂNTICA

Com o intuito de caracterizar o notável crescimento das pesquisas sobre a Web Semântica no cenário mundial nos últimos dez anos, foi realizado um levantamento bibliométrico sobre o tema nas duas principais bases de dados de indexação de publicações científicas, a *Web of Science* e a *Scopus*, de forma a obter um panorama geral destas pesquisas por ano e por área de conhecimento em que foram realizadas e publicadas.

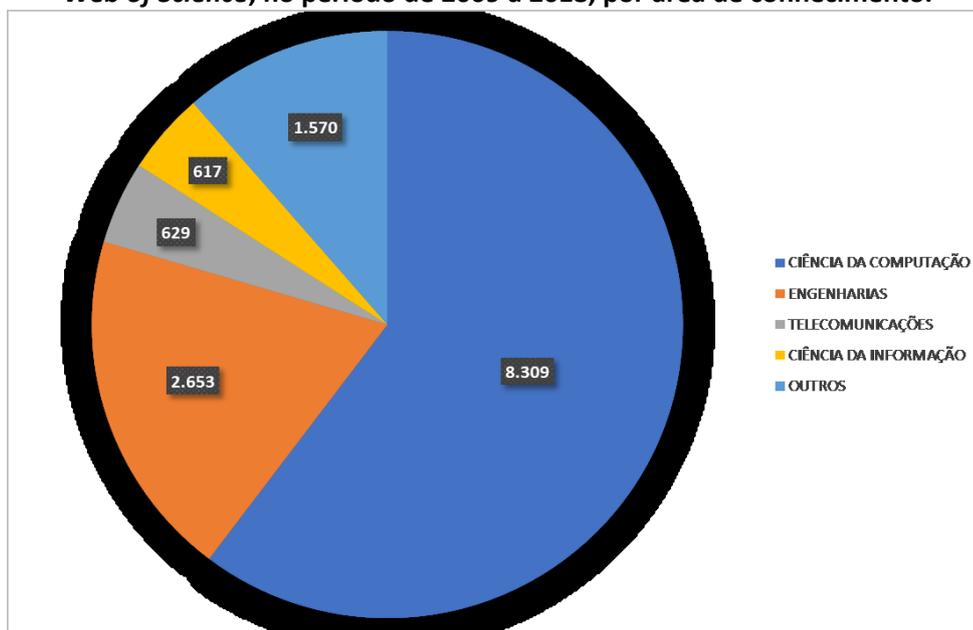
Inicialmente foi realizada uma busca na coleção principal da base de dados *Web of Science*, onde a chave de busca utilizada foi [“semantic web” OR “linked open data” OR “web of data”], dentro do campo tópico, no período de 2009 a 2018, ou seja, abrangendo os últimos dez anos.

Como resultado desta busca foram encontrados 10.722 artigos publicados sobre estes assuntos no período considerado. Este resultado foi então ordenado por área de conhecimento do periódico onde cada artigo foi publicado, possibilitando a elaboração do gráfico apresentado na Figura 5.

É importante frisar que um periódico pode estar cadastrado em mais de uma área de

conhecimento nesta base de dados, o que faz com que a somatória dos valores apresentados no gráfico seja ligeiramente superior ao número de resultados retornados pela busca.

Figura 5: Trabalhos sobre Web Semântica publicados em periódicos indexados na *Web of Science*, no período de 2009 a 2018, por área de conhecimento.



Fonte: Elaborado pelos autores, com dados extraídos da base de dados *Web of Science* (Clarivate Analytics).

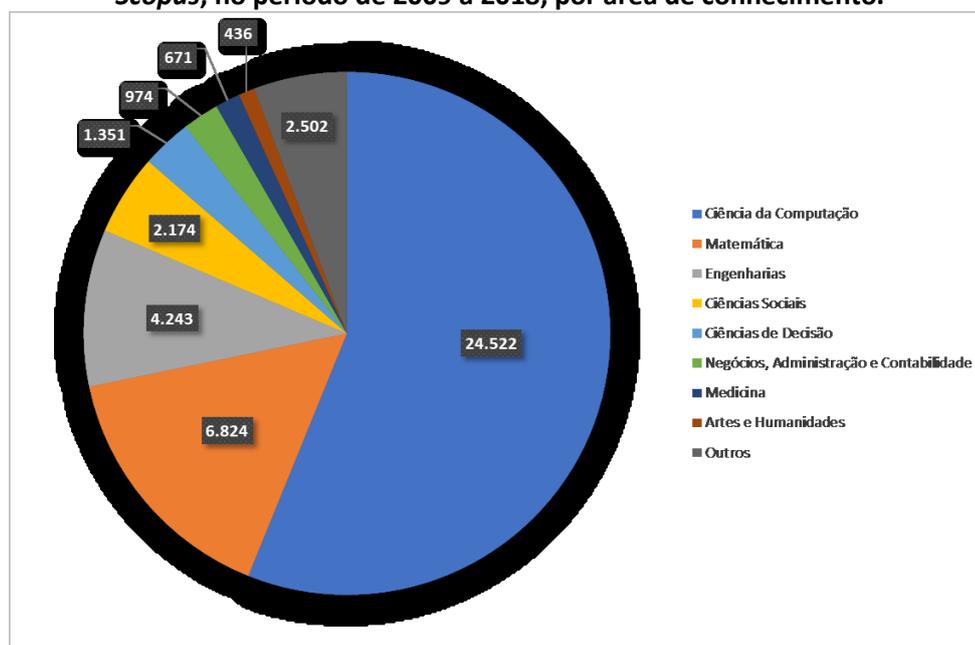
A partir dos valores apresentados no gráfico da Figura 5, é possível identificar que a área da ciência da computação originou mais da metade dos trabalhos publicados sobre o tema Web Semântica, seguida pelas áreas de engenharia e telecomunicações.

Ainda de acordo com o gráfico, a área da Ciência da Informação teve publicados apenas 617 trabalhos sobre este mesmo tema nos últimos dez anos, o que representa menos de 5% do que as outras áreas já mencionadas publicaram no mesmo período.

Em seguida, outra busca foi realizada na base de dados *Scopus* utilizando os mesmos termos [“semantic web” OR “linked open data” OR “web of data”] e o mesmo período (2009 a 2018) utilizados na busca anterior.

Como resultado foram encontrados 28.187 artigos, o qual também foi ordenado por área de conhecimento, o que possibilitou a elaboração do gráfico apresentado na Figura 6. Da mesma forma, como na busca anterior, é importante frisar que, também nesta base, um periódico pode estar cadastrado em mais de uma área de conhecimento, o que faz com que a somatória dos valores apresentados no referido gráfico seja ligeiramente superior ao número de resultados retornados pela busca.

Figura 6: Trabalhos sobre Web Semântica publicados em periódicos indexados na *Scopus*, no período de 2009 a 2018, por área de conhecimento.



Fonte: Elaborado pelos autores, com dados extraídos da base de dados *Scopus* (Elsevier).

Os resultados obtidos na busca realizada na base de dados *Scopus* também apontam que a maior parte dos trabalhos publicados sobre o tema Web Semântica, em periódicos indexados nesta base, são oriundos da área da Ciência da Computação (mais da metade), da matemática e das engenharias. Na *Scopus*, os periódicos da área de Ciência da Informação estão agrupados dentro da área de conhecimento Artes e Humanidades, na qual foram encontrados 436 artigos sobre o tema procurado, o que representa menos de 2% do que foi publicado pelas outras áreas mencionadas.

3.1 EVOLUÇÃO DAS PESQUISAS SOBRE WEB SEMÂNTICA NA CIÊNCIA DA INFORMAÇÃO

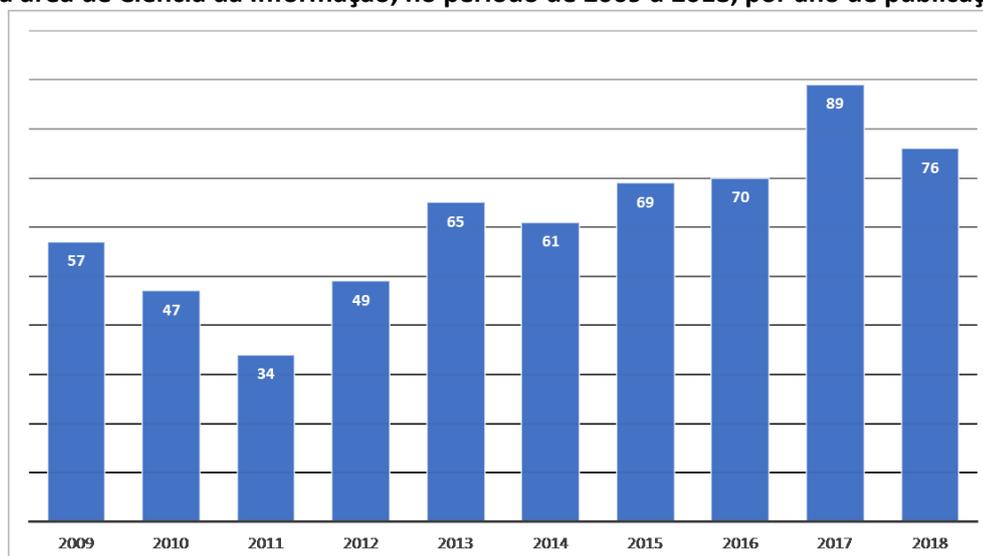
Numa segunda etapa deste trabalho, a partir dos resultados das buscas realizadas nas bases de dados *Web of Science* e *Scopus* sobre o tema Web Semântica, foram filtrados somente os trabalhos publicados, para o mesmo período (2009-2018), em periódicos da área de Ciência da Informação (*Information Science Library Science*), na base de dados *Web of Science*, e da área Artes e Humanidades (*Arts and Humanities*) na base de dados *Scopus*, os quais foram considerados nesta análise.

Primeiramente, os 617 trabalhos sobre o tema Web Semântica encontrados na busca realizada na *Web of Science*, que foram publicados no período 2009 a 2018 em periódicos

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

classificados como da área de Ciência da Informação, foram agrupados por ano de publicação, o que possibilitou a elaboração do gráfico apresentado na Figura 7.

Figura 7: Trabalhos sobre Web Semântica publicados em periódicos indexados na *Web of Science*, na área de Ciência da Informação, no período de 2009 a 2018, por ano de publicação.

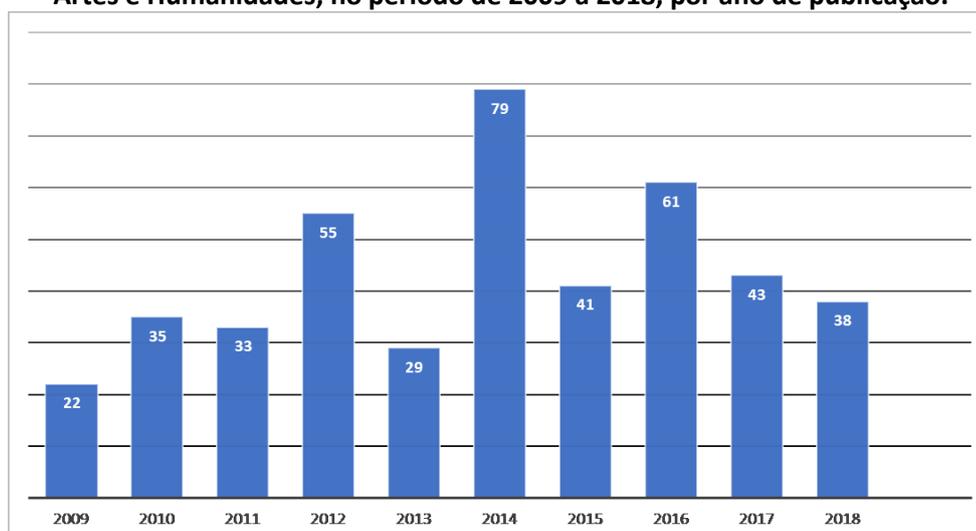


Fonte: Elaborado pelos autores, com dados extraídos da base de dados *Web of Science* (Clarivate Analytics).

Os valores apresentados no gráfico da Figura 7 mostram que, a partir de 2013, houve um expressivo aumento na quantidade de trabalhos publicados sobre o tema Web Semântica dentro da área da Ciência da Informação, o que se deu de forma mais acentuada nos últimos dois anos (2017-2018). Isso mostra que pode ter havido um maior envolvimento desta área, e por conseguinte do profissional da informação, nas pesquisas sobre o tema neste período.

Em seguida, os 436 trabalhos sobre o tema Web Semântica encontrados na busca realizada na *Scopus*, que foram publicados no mesmo período em periódicos classificados como da área Artes e Humanidades, foram também agrupados por ano de publicação, o que possibilitou a elaboração do gráfico apresentado na Figura 8.

Figura 8: Trabalhos sobre Web Semântica publicados em periódicos indexados na *Scopus*, na área Artes e Humanidades, no período de 2009 a 2018, por ano de publicação.

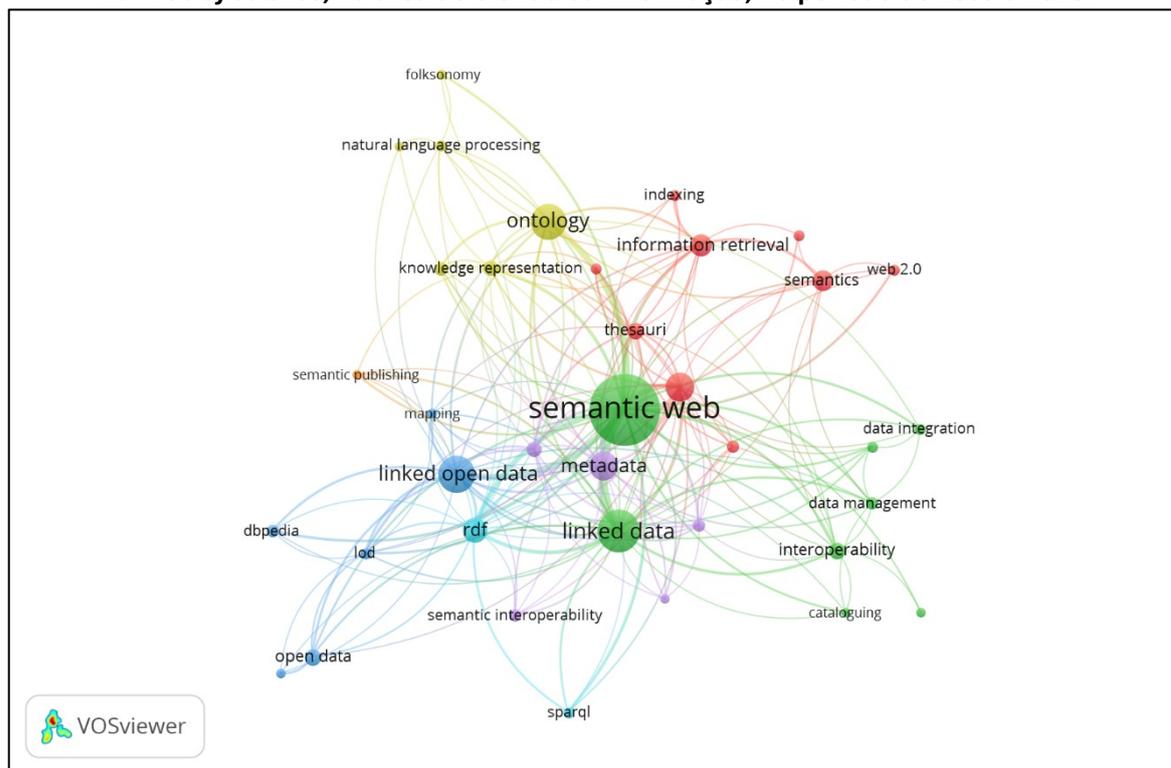


Fonte: Elaborado pelos autores, com dados extraídos da base de dados *Scopus* (Elsevier).

Neste caso foram registrados picos de publicações sobre o tema, dentro da área de Ciência da Informação, nos anos de 2012, 2014 e 2016. Os dois últimos anos deste período mostram uma tendência de retração na quantidade de publicações por esta área, porém os números ainda são maiores do que a média registrada nos primeiros cinco anos do período analisado (2009-2013).

Por fim, os metadados dos trabalhos filtrados, tanto na busca na *Web of Science* quanto na busca na *Scopus*, foram exportados em arquivos no formato texto, os quais em seguida foram importados no software *VOSviewer* com o propósito de efetuar uma análise de coocorrência de palavras-chave informadas pelos respectivos autores destes trabalhos. A partir das análises efetuadas por meio do referido software foi possível obter os mapas de coocorrência de palavras-chave ilustrados na Figura 9 e na Figura 10 a seguir. Isto permite verificar quais os assuntos mais abordados pelos trabalhos publicados sobre Web Semântica na área da Ciência da Informação.

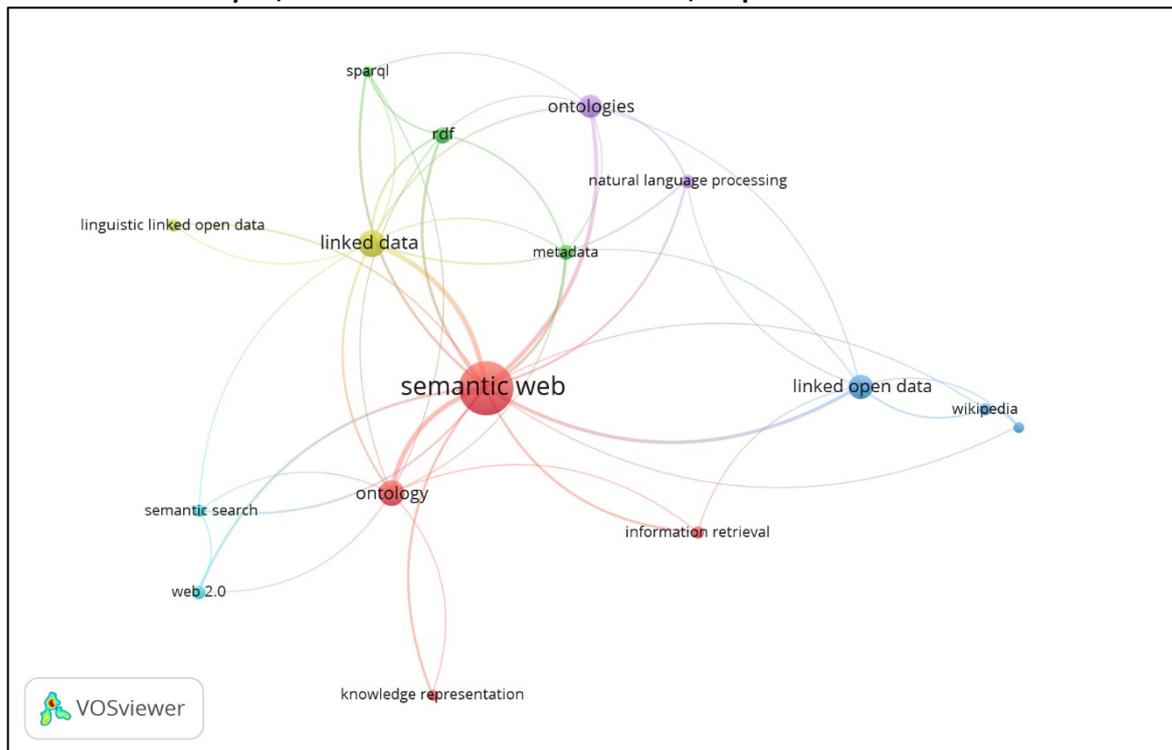
Figura 9: Mapa de coocorrência de palavras-chave dos trabalhos sobre Web Semântica indexados na *Web of Science*, na área de Ciência da Informação, no período de 2009 a 2018.



Fonte: Software *VOSviewer*, com dados extraídos da base de dados *Web of Science* (*Clarivate Analytics*).

O mapa ilustrado na Figura 9 mostra que os trabalhos sobre Web Semântica indexados na *Web of Science* (na área de Ciência da Informação) abordam predominantemente, de acordo com as palavras-chave informadas pelos autores, assuntos como dados ligados, dados abertos ligados, metadados, ontologias, modelo RDF, recuperação da informação, entre outros. Observar no mapa que, quanto maior o tamanho do nó maior a quantidade de trabalhos nos quais a respectiva palavra-chave foi mencionada, e quanto maior a espessura do arco que liga dois nós maior a coocorrência das respectivas palavras-chave em um mesmo trabalho.

Figura 10: Mapa de coocorrência de palavras-chave dos trabalhos sobre Web Semântica indexados na *Scopus*, na área de Artes e Humanidades, no período de 2009 a 2018.



Fonte: Software *VOSviewer*, com dados extraídos da base de dados *Scopus* (*Elsevier*).

Por sua vez, o mapa ilustrado na Figura 10 mostra que os trabalhos sobre Web Semântica indexados na *Scopus* (na área de Artes e Humanidades), apesar de apresentarem uma quantidade menor de palavras-chave, abordam predominantemente os mesmos assuntos quando comparado aos trabalhos indexados na *Web of Science*. Observar neste mapa que apesar da menor quantidade de nós, aqueles de maior tamanho são praticamente os mesmos do mapa da Figura 9, assim como as espessuras dos arcos, o que indicam as mesmas tendências de coocorrência das respectivas palavras-chave em um mesmo trabalho.

3.2 ANÁLISE DOS RESULTADOS

Os resultados apresentados na forma de gráficos nas Figuras 1 a 4, permitem deduzir que a quantidade de pesquisas sobre o tema Web Semântica, dentro da área da Ciência da Informação, ainda é pouco expressiva quando se comparado às pesquisas realizadas sobre o mesmo tema nas áreas da Ciência da Computação e das Engenharias, por exemplo.

Mas, apesar de ser uma quantidade ainda pouco expressiva frente às demais áreas que realizam pesquisas sobre o tema, é possível afirmar que tem havido um crescente interesse pelo tema dentro da Ciência da Informação nos últimos cinco anos (2014 a 2018). As

mudanças no perfil do profissional da informação que tem ocorrido ao longo das últimas décadas, para acompanhar as transformações na forma como passamos a consumir informação neste mundo cada vez mais digital e orientado a dados (estamos na era da ciência de dados), bem como o surgimento de linhas de pesquisas voltadas a temas como informação e tecnologia, ambientes informacionais digitais, curadoria digital, evolução da Web, entre outras, podem ser considerados como o motor de indução deste crescente interesse pelo tema na Ciência da Informação.

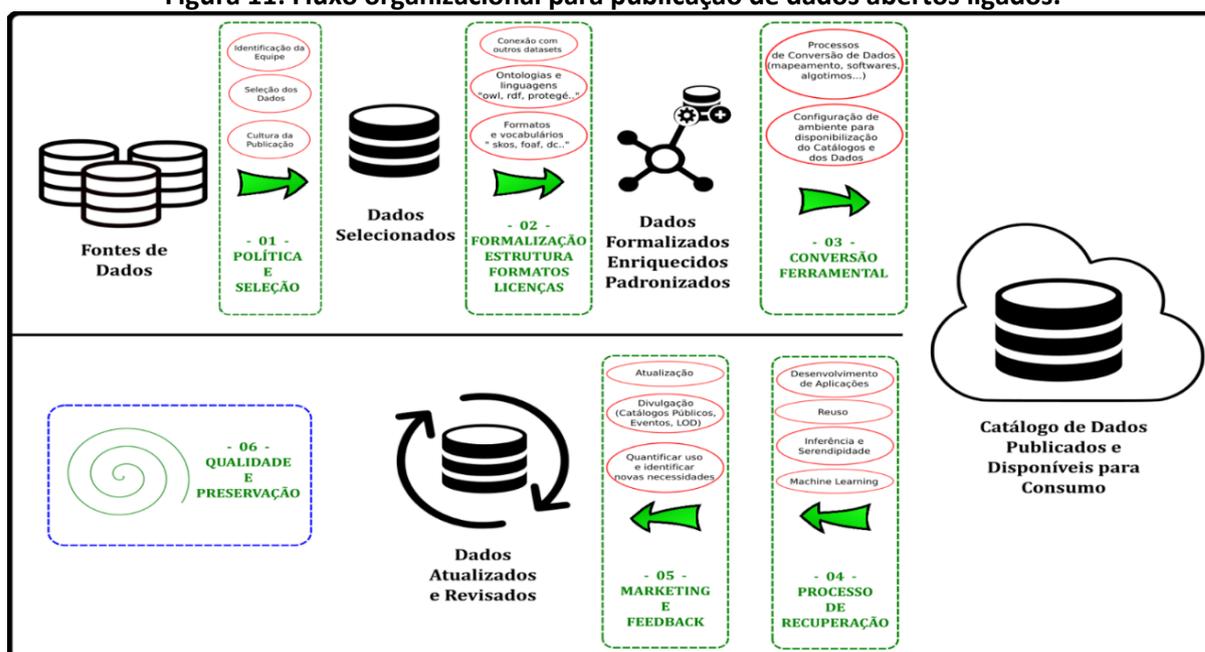
A análise de coocorrência das palavras-chave dos trabalhos publicados na área da Ciência da Informação, apresentada nas figuras 9 e 10, permitiram constatar que os assuntos mais frequentemente abordados nestes trabalhos, dentro do tema Web Semântica, foram dados ligados, dados abertos ligados, metadados, ontologias, modelo RDF e recuperação da informação. Isto permite estabelecer um perfil e perspectivas de atuação do profissional da informação nas atividades de pesquisas sobre o tema.

Por outro lado, a crescente demanda por profissionais da informação para atuarem em atividades desenvolvimento relacionados à Web Semântica, e a busca pelo entendimento e aperfeiçoamento dos modelos e *frameworks* desenvolvidos nas áreas da Ciência da Computação e das Engenharias, por exemplo, ampliam as perspectivas de atuação do profissional da informação neste campo, como será visto a seguir, o que pode trazer enormes contribuições para a evolução e consolidação da Web Semântica.

4 PERSPECTIVAS PARA O PROFISSIONAL DA INFORMAÇÃO NA PESQUISA E DESENVOLVIMENTO EM WEB SEMÂNTICA

Conforme já mencionado anteriormente, o desenvolvimento e sucesso da Web Semântica dependem da criação e publicação massiva de dados abertos ligados, seguindo as boas práticas do Linked Open Data. Diante disso, e levando-se em consideração o **fluxo para publicação de dados abertos ligados** proposto por Santarem Segundo (2018), ilustrado na Figura 11, é possível identificar algumas atividades e processos que podem ser desempenhados de forma mais eficaz por profissionais da informação.

Figura 11: Fluxo organizacional para publicação de dados abertos ligados.



Fonte: Santarem Segundo (2018).

A publicação de dados abertos ligados e o desenvolvimento de aplicações para o consumo destes dados abrem grandes perspectivas de atuação para o profissional da informação. Da mesma forma, o surgimento de linhas de pesquisas dentro da Ciência da Informação voltadas para o estudo da informação digital, tecnologia e ciência de dados, ampliam ainda mais as perspectivas de atuação do profissional da informação da era atual, desde que possuam as competências necessárias para tais atividades.

Neste contexto, competência é definida como o conjunto de qualificações (conhecimentos, habilidades e atitudes) que credenciam um profissional a exercer determinado trabalho (CARBONE *et al.*, 2009). Na literatura, os trabalhos publicados por Raju (2017) e de Vallejo-Echavarría (2017) se propuseram a mapear algumas competências do moderno profissional da informação para atuar em ambientes digitais de informação, as quais foram tomadas como base para a elaboração do Quadro 1, que apresenta uma lista de atividades relacionadas à pesquisa e desenvolvimento em Web Semântica, definidas com base no fluxo de atividades ilustrado na Figura 11, para as quais o profissional da informação reúne competências para atuar e se sobressair frente aos profissionais de outras áreas de conhecimento.

A lista de atividades elencadas no Quadro 1 não é exaustiva, uma vez que a proposta do presente trabalho é apenas e tão somente a prospecção de perspectivas de atuação para

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

o profissional da informação, a partir das competências que possui para o desenvolvimento das atividades propostas sob o enfoque da teoria e preceitos da Ciência da Informação, de forma a aperfeiçoar e dar aplicabilidade ao desenvolvimento efetuado em outras áreas do conhecimento, principalmente a Ciência da Computação e as Engenharias.

Quadro 1: Atividades relacionadas à Web Semântica nas quais o profissional da informação reúne competências para se sobressair em relação a profissionais de outras áreas do conhecimento.

Atividade	Competências do Profissional da Informação para atuar na atividade
Formulação de políticas e promoção à cultura de publicação de dados	<ul style="list-style-type: none"> ● Atuar na criação, organização e preservação de conteúdos digitais;
Seleção de dados para publicação	<ul style="list-style-type: none"> ● Selecionar, registrar, armazenar, recuperar e difundir a informação gravada em qualquer meio para os usuários de unidades, serviços e sistemas de Informação.
Definição de formatos e licenças e construção de vocabulários controlados e ontologias	<ul style="list-style-type: none"> ● Descrever recursos de informação em ambientes digitais; ● Definir regras léxicas e lógica para construção de vocabulários controlados e ontologias;
Definição e monitoramento dos processos de conversão ferramental de dados	<ul style="list-style-type: none"> ● Articular, junto com profissionais da área de computação, a integração coerente e consistente de ferramentas para a operação de serviços em sítios Web.
Criação e atualização de catálogos de dados	<ul style="list-style-type: none"> ● Descrever recursos de informação em ambientes digitais; ● Atuar na criação, organização e preservação de conteúdos digitais.
Recuperação e reuso da informação na Web Semântica	<ul style="list-style-type: none"> ● Selecionar, registrar, armazenar, recuperar e difundir a informação gravada em qualquer meio para os usuários de unidades, serviços e sistemas de Informação;
Apoio ao desenvolvimento e <i>User Experience</i> em aplicativos e sítios para a Web Semântica	<ul style="list-style-type: none"> ● Desenhar modelos de organização e esquemas de comunicação, de navegação e de metadados em sítios Web; ● Desenhar e pensar funcionalidades que estejam alinhadas aos objetivos a serem alcançados, de maneira a gerar experiências inovadoras.
Atualização e divulgação dos catálogos de dados	<ul style="list-style-type: none"> ● Selecionar, registrar, armazenar, recuperar e difundir a informação gravada em qualquer meio para os usuários de unidades, serviços e sistemas de Informação;
Identificação e especificação de necessidades dos usuários	<ul style="list-style-type: none"> ● Planejar e executar estudos de usuários de sítios Web e sistemas de informação; ● Recomendar estruturas de serviços em sítios Web.
Garantia da qualidade dos dados publicados	<ul style="list-style-type: none"> ● Desenhar esquemas de metadados e estratégias de codificação de dados.
Preservação dos dados publicados	<ul style="list-style-type: none"> ● Atuar na criação, organização e preservação de conteúdos digitais;

Fonte: Elaborado pelos autores.

Com a evolução dos estudos relacionados a Web Semântica na área da Ciência da Informação, bem como o interesse em gerar essa nova perspectiva, o Quadro 1 apresentado poderá sofrer mudanças. Como ressaltado, as competências dos Profissionais da Informação

relatados neste trabalho são baseados nos estudos apresentados nestes últimos 5 anos de interesse pelo tema, verificando assim essas possíveis atribuições de competências.

Hendler (2003) já afirmava que, para que a Web Semântica venha realmente a funcionar, há a necessidade das áreas de conhecimento se complementarem, ou seja, ela não alcançará sua completude se o conteúdo (metadados) e as ferramentas não forem amplamente compartilhados. Sabemos que a World Wide Web se originou pela expansão em massa de código aberto, do software livre, e a disseminação de tecnologias de computação de baixo custo. Para a Web Semântica não será diferente, irá requerer toda essa disseminação, para que esta realmente venha revolucionar o conhecimento humano e as máquinas.

5 CONSIDERAÇÕES FINAIS

As tecnologias e ferramentas relacionadas à Web Semântica representam as principais inovações nesta era da ciência de dados na qual vivemos atualmente. Neste sentido, este trabalho contextualizou que estamos passando por uma transição da atual Web de documentos (consumível por humanos) para uma Web de dados (consumível por humanos e máquinas), e trouxe uma revisão do que é a Web Semântica e quais suas ferramentas, padrões e formatos.

A partir de um levantamento bibliométrico foi possível apontar que as pesquisas relacionadas à Web Semântica estão em crescente evolução nos últimos dez anos, porém ainda fortemente concentradas na Ciência da Computação e nas Engenharias. Entretanto, as pesquisas sobre este tema na área da Ciência da Informação, apesar de ainda pouco expressivas em termos de quantidade quando comparado às outras áreas mencionadas, apresentam tendência de crescimento nestes últimos cinco anos, o que abre perspectivas de atuação para os profissionais da informação nestas pesquisas, a partir do surgimento de linhas de pesquisas dentro da Ciência da Informação voltadas para o estudo da informação digital, tecnologia e ciência de dados. Além disso, foram identificados os assuntos mais abordados nas pesquisas nesta área a partir da análise de coocorrência de palavras-chaves.

Além disso, as atividades e processos de publicação de dados abertos ligados e de desenvolvimento de aplicações para a Web Semântica também podem oferecer, num futuro próximo, boas perspectivas de atuação para o profissional da informação, o que certamente contribuirá para o avanço e consolidação destas tecnologias.

**XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC**

Neste sentido, o trabalho apresentou uma tabela que propõe uma série de atividades nas quais o profissional da informação pode se sobressair frente à profissionais de outras áreas do conhecimento, dadas as competências adquiridas durante sua formação.

Por fim, como conclusão deste trabalho, espera-se que ele sirva de incentivo para os atuais e futuros profissionais da informação voltarem suas atenções e perspectivas profissionais para contribuírem com a consolidação da Web Semântica e, conseqüentemente, da Web de dados, revolucionando a forma como consumimos informação nesta era da ciência de dados.

Como trabalhos futuros propõe-se um estudo detalhado dos assuntos relacionados à Web Semântica mais abordados nas pesquisas realizadas dentro da Ciência da Informação e suas contribuições para as pesquisas realizadas em outras áreas do conhecimento, como a Ciência da Computação e as Engenharias. Nesta mesma linha também é proposta a realização de um levantamento detalhado das competências relacionadas à tecnologia e ciência de dados adquiridas pelos profissionais da informação durante sua formação, em nível de graduação e pós-graduação, e quais as atividades ou funções nestas áreas que tem demandado este tipo de profissional.

REFERÊNCIAS

ALLEMANG, D.; HENDLER, J. A. **Semantic Web for the working ontologist: modeling in RDF, RDFS and OWL**. BURLINGTON: Morgan Kaufmann, 2008. 330 p.

AMARAL, F. **Introdução a Ciência de Dados: mineração de dados e Big Data**. Rio de Janeiro: Alta Books, 2016. 320 p.

ANTONIOU, G.; HARMELEN, F. **A Semantic Web primer**. 2ed. Cambridge: MIT Press, 2008. 264 p.

ARTHUR, Charles. Google's Marissa Mayer on the importance of real-time search. **The Guardian**, Londres, 8 jul. 2009. Disponível em: <<https://www.theguardian.com/technology/2009/jul/08/google-search-marissa-mayer>>. Acesso em: 16 jul. 2019.

BERNERS-LEE, T.; HENDLER, J.; LASSILA, O. The Semantic Web: a new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. **Scientific American**, New York, v. 284, 2001, p. 28-37.

CARBONE, P. P. et al. **Gestão por competências e gestão do conhecimento**. 3ed. Rio de Janeiro: Editora FGV, 2009.

XX ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO – ENANCIB 2019
21 a 25 de outubro de 2019 – Florianópolis – SC

CARDOSO, J. **Semantic Web services: theory, tools and applications**. Hershey: IGI Global, 2007. 353 p.

EIS, D. **Introdução à Web Semântica: A Inteligência da Informação**. São Paulo: Casa do Código, 2016. 170 p.

HENDLER, J. Science and the Semantic Web. **Science**, Washington, Vol. 299, no. 5606, 2003, p. 520-521. DOI: <https://doi.org/10.1126/science.1078874>. Disponível em: <<https://science.sciencemag.org/content/299/5606/520>>. Acesso em 16 jul. 2019.

HERT, M. et al. UpLink: a linked data editor for BDR-to-RDF data. In: INTERNATIONAL CONFERENCE ON SEMANTIC SYSTEMS, 7., 2011, New York. **Proceedings...** New York: ACM, 2011. p. 159-162. Disponível em: <<http://doi.acm.org/10.1145/2063518.2063539>>. Acesso em: 20 set. 2012.

LAUFER, C. **Guia de Web Semântica**. São Paulo: Projeto SPUK, 2015. Disponível em: <https://nic.br/media/docs/publicacoes/13/Guia_Web_Semantica.pdf>. Acesso em: 23 nov. 2017.

LING, H.; ZHOU, S. Translating relational databases into RDF. In: INTERNATIONAL CONFERENCE ON ENVIRONMENTAL SCIENCE AND INFORMATION APPLICATION TECHNOLOGY, 2., 2010, Wuhan. **Proceedings...** Piscataway: IEEE, 2010. p. 464-467.

LV, Y.; MA, Z. M. Transformation of relational model to RDF model. In: IEEE INTERNATIONAL CONFERENCE ON SYSTEMS, MAN AND CYBERNETICS, 2008, Singapore. **Proceedings...** Piscataway: IEEE, 2008. p. 506-511.

MACHADO, A. L.; PARENTE DE OLIVEIRA, J. M. DIGO: an open data architecture for e-government. In: ENTERPRISE DISTRIBUTED OBJECT COMPUTING CONFERENCE WORKSHOPS, 15., 2011, Helsinki. **Proceedings...** Piscataway: IEEE, 2011. p. 448-456.

MOHAMED, H.; JINCAI, Y.; QIAN, J. Towards integration rules of mapping from relational databases to Semantic Web ontology. In: INTERNATIONAL CONFERENCE ON WEB INFORMATION SYSTEMS AND MINING, 2010, Sanya. **Proceedings...** Piscataway: IEEE, 2010. p. 335-339.

RAJU, J. Information Professional or IT Professional? The Knowledge and Skills Required by Academic Librarians in the Digital Library Environment. **Libraries and the Academy**, vol. 17 no. 4, 2017, p. 739-757. DOI: <https://doi.org/10.1353/pla.2017.0044>.

SANTAREM SEGUNDO, J. E. Web Semântica: fluxo para publicação de dados abertos e ligados. **Informação em Pauta**, Fortaleza, v. 3, número especial, 2018, p. 117-140. DOI: <https://doi.org/10.32810/2525-3468.ip.v3iEspecial.2018.39721.117-140>.

VALLEJO-ECHAVARRÍA, J. C. El rol del profesional de la información en escenarios digitales [Entrevista con Diego Ferreyra, marzo 15 de 2016]. **Revista Interamericana de Bibliotecología**, vol. 40(1), 2017, p. 87-92. DOI: <https://doi.org/10.17533/udea.rib.v40n1a08>.