

Criação de padrões na web semântica: perspectivas e desafios

Maria José Vicentini Jorente

Doutora; Universidade Estadual Paulista Júlio de Mesquita Filho, São Paulo, SP, Brasil;
mjjorente@yahoo.com

Mariana Cantisani Padua

Doutoranda; Universidade Estadual Paulista Júlio de Mesquita Filho, São Paulo, SP, Brasil;
mariana.cantisani@gmail.com

José Eduardo Santarem Segundo

Doutor; Universidade de São Paulo, Ribeirão Preto, SP, Brasil;
santarem@usp.br

Resumo: Aborda a criação de padrões de ontologias e de que forma os recursos e práticas desenvolvidas pela equipe do Ontology Design Patterns (Projeto NEON) propõem normas e recomendações visando a organização, o acesso e a recuperação de recursos na Web Semântica. Pensar a complexidade na criação dos padrões de linguagens baseados em ontologias implica entender que ela se expressa como uma inquietude, ou seja, como resolver problemas surgidos da convergência dos sistemas implicados e integrados na Web. Desse ponto de vista, existem diversos padrões e modelos para a criação de *design patterns*, baseados em padrões de conteúdo, de conhecimento, linguísticos, lógicos, entre outros. O problema tratado nesse artigo é como um padrão de *design* provindo da ODP facilita práticas informacionais de indexação e recuperação da informação. A pesquisa, exploratória e descritiva, objetivou elucidar padrões e normas de *design* de ontologias, fornecendo critérios e métodos a fim de formular soluções para a criação dos padrões de ontologias facilitadores da recuperação e acesso à informação, por meio de uma Web semântica. As propostas do artigo justificam-se pela multiplicação e convergência dos meios e suportes de informação em plataformas e ambientes digitais, verificando-se um aumento dos dados e provocando urgente necessidade de gestão desses dados a nível mundial. Conclui que a investigação dessa temática em Ciência da Informação (CI) é importante para uma formação dos profissionais da informação, pois as dificuldades de recuperação da informação podem ser dirimidas por uma CI que explore os recursos de *design* de ontologias.

Palavras-chave: *Onthology Design Patterns*. Web semântica. Sistema Complexo. Convergências.

1 Introdução

O conceito de cultura como a soma de vestígios de memórias embutidos nas linguagens do cotidiano (MOLES, 1974, p. 13); como a soma de fragmentos da informação circulante no cotidiano recolhidos em função das necessidades imediatas, preenchidos e re-preenchidos por significados nos processos de utilização de conceitos, das linguagens e dos aparatos-suporte de informação, ou ainda, como “cultura mosaico” tem sido, desde a sua criação, uma definição muito adequada para fundamentar os processos e os produtos culturais da modernidade tardia e pós-modernidade, dado o seu caráter complexo e fragmentário. Pensar a complexidade, assim constituída, implica entender que ela se expressa como uma inquietude e como uma pergunta autorreferente quanto a sua natureza. Para respondê-la é necessário entender que a principal qualidade do que é complexo é conjugar muitos elementos ou partes, o que pode ser observável sob vários aspectos. Assim, na Ciência, a complexidade encerra a mudança da relação entre o todo e as suas partes.

Na contemporaneidade, de acordo com esse princípio, a complexidade sistêmica do processo informacional e comunicacional relaciona elementos e partes. Da convergência das informações nas estruturas sistêmicas complexas, emergem novas formas representacionais e de apresentação da informação. Formas hipertextuais (complexas) de apresentação da informação interferiram nas interações dos indivíduos com a informação e criaram novas cognições e um *redesenho* (reapresentação) da Cultura: as novas formas imagético-textuais, criadas pelo hipertexto, proporcionaram uma nova forma de distribuição da inteligência e da cognição, afetando a maneira como lemos e, também, como escrevemos, borrando não só as fronteiras entre leitura e escritura, bem como de autoria, a partir das possibilidades oferecidas de edição e colaboração. O hipertexto transforma cognitivamente as pessoas que interagem como parte de um mesmo processo e fazem com que a escrita seja uma tarefa menos individual e mais coletiva e colaborativa.

As representações modificadas pela estética, veiculada nas mídias da Web e usadas como meio de disseminação de informação e de conhecimento de muitos para muitos, criaram novos planos de contato na rede de percepções e na

cognição dos agentes, sejam humanos ou maquínicos. Como afirma o catedrático do Departamento de Engenharia de Sistemas da Universidade Politécnica de Madrid, especializado em arquitetura e tecnologia de computadores, Fernando Sáez Vacas (2006, p. 2, tradução nossa),

O que é evidente e inevitável é que se está produzindo um processo acelerado de impregnação tecnológica de nossas vidas, em que a grande e crescente complexidade da tecnologia e a ainda maior complexidade das relações humanas com ela requerem respostas muito pensadas da índole tecno-cultural, o que não está se dando.

Sáez Vacas (2006) afirma, também, que ainda não estamos preparados para usar com critério nem eficácia a funcionalidade das tecnologias que passam por nossas mãos no cotidiano, de acordo com o autor existe uma necessidade de compreender ou administrar o Novo Entorno Tecnosocial (NET)

[...] e dele extrair o melhor de suas extraordinárias oportunidades informativas, de comunicação, artísticas, científicas, de entretenimento, sanitárias e operativas de muitas espécies, também educativas mas sem renunciar ao que mereça conservar-se (SÁEZ VACAS, 2007, p. 2, tradução nossa).

Embora, contemporaneamente, seja possível verificar um aumento significativo de informações disponibilizadas na Web (WWW, *World Wide Web*), na qual a quantidade de dados cresce exponencialmente. O mesmo autor enfatizava, anteriormente, em 2004, que há um conceito-chave - a co-evolução humano-tecnologia - esquecido ou simplesmente desconhecido devido ao crescente grau de especialização das sociedades em que se presentifica o fenômeno da exponencialidade do crescimento. Tal movimento de coevolução no relacionamento humano-tecnologia, para os autores do artigo, dá-se pelos códigos e pelas linguagens, por meio das quais se possibilita o relacionamento. Linguagens, por sua vez, apresentam uma sintaxe e uma decorrente semântica que da primeira deriva de maneira complexa.

Para estudar a complexidade da criação exponencial da informação em rede, seguindo o princípio do relacionamento via linguagens, Sáez Vacas criou, em 2007, uma metodologia de análise de 21 dimensões distribuídas em cinco categorias. Uma dessas categorias, *Transformações na representação, na*

captação e no manejo de informação, nos interessa, de maneira especial, no presente estudo e divide-se em quatro dimensões distintas:

- a) potencialidade - as infotecnologias mostram-se como ambientes para a construção de uma realidade que nunca se mostra realizada por completo. Com elas se tem a ampliação constante da capacidade de armazenamento e da velocidade dos processadores, o crescimento das redes de comunicação, o surgimento de novas funcionalidades e a implementação de serviços inovadores;
- b) oniprocessabilidade - capacidade de processamento estendido por todos os nós da infraestrutura: telefones móveis, pontos de acesso *wireless*, sensores, implantes médicos, etc.;
- c) análogo-digitalidade - a fronteira entre os mundos real e digital está cada vez mais imperceptível, especialmente, para os nativos digitais;
- d) digitalidade - o alfabeto binário, construído somente por dois símbolos (0 e 1) é a verdadeira linguagem universal das máquinas que forma a trama infotecnológica da rede.

É essa última dimensão que trata do aspecto sintático da linguagem universal das máquinas e que forma a trama infotecnológica da rede complexa de informação e sua interação com a cultura contemporânea. Elegeu-se a linguagem binária para ser discutida no seu patamar semântico, pois para se ter uma noção da dimensão da importância dessa problemática, destaca-se que, em 2014, foram contabilizados 1014 conjuntos de dados, de acordo com a publicação *State of the LOD cloud* do Linked Data (SCHMACHTENBERG; BIZER; PAULHEIM, 2014). As tecnologias *Linked Data* estão sendo utilizadas para compartilhar dados que abrangem uma vasta gama de domínios de diferentes tópicos. O Quadro 1 dá uma visão geral dos domínios por tópicos.

Quadro 1 – Visão geral dos conjuntos de dados

Tópicos	Conjunto de dados	Porcentagem
Governos	183	18.05%
Publicações	96	9.47%
Ciências da vida	83	8.19%
Conteúdo gerado por usuários	48	4.73%
Meta-domínio	41	4.04%
Media	22	2.17%
Geográficos	21	2.07%

Web Social	520	51.28%
Total	1014	

Fonte: Schmachtenberg, Bizer e Paulheim (2014).

De acordo com Heath e Bizer (2011), o termo *Linked Data* refere-se a um conjunto de melhores práticas para publicação e conexão de dados estruturados na Web, permitindo estabelecer ligações entre itens de diferentes fontes de dados para formar um único espaço de dados global. Para compreender esse universo apresenta-se o conceito de *Linked Data*: dados ligados entre si, relacionados de acordo com um conjunto de práticas introduzidas por Berners-Lee (2001), com a função de estruturar, organizar e publicar dados na Web.

Meios e suportes de informação multiplicam-se e convergem em plataformas e ambientes digitais, verificando-se um aumento metastático dos dados produzidos e provocando a urgente necessidade de gestão profunda e ampla destes dados a nível mundial. Há, para isso, cuidados que necessitam ser tomados a fim de que o armazenamento desta quantidade de dados possa ser gerenciado e a fim de prover uma estabilidade, dada a quantidade de acessos e os diversos tipos de usuários. Tais cuidados devem possibilitar ao sistema retornar uma busca satisfatória e precisa, tornando o processo de recuperação da informação onipresente e expressivo.

Berners-Lee (2001), em seu texto introdutório sobre a Web Semântica, afirma que a principal função de se estruturar os dados é para que a recuperação da informação ocorra de forma dinâmica e simples, e, por esta razão, é necessário que ocorra a padronização da utilização das ontologias que são empregadas nos diversos domínios existentes. Ainda, segundo o autor, a tecnologia já existe, porém é necessário solucionar uma série de problemas relacionados ao desenvolvimento das bases de dados. É, por exemplo, na fase inicial da construção de uma ontologia, que a utilização das terminologias apropriadas garante que as informações possam ser inseridas de forma correta, para que a interoperabilidade e a integração entre os bancos de dados ocorra de forma coerente e bem definida.

Neste contexto, uma das principais questões levantadas refere-se às formas de definição de ontologias ou de vocabulários controlados, para as mais

diversas áreas do conhecimento, permitindo representar e relacionar, explicitamente, a semântica dos dados. Nesse campo, no qual muitos pontos estão sendo levantados, é importante ressaltar que as relações entre os conteúdos de um recurso (qualquer objeto que pode ser atingido através de uma URI – *Uniform Resource Identifier*) devem relacionar-se com os demais recursos na Web. É essencial que esses sejam suficientemente expressivos, para que as máquinas ou agentes sejam capazes de entender o real significado do dado e para que possam ser relacionados, intermediando as necessidades de cada usuário e, assim, apresente as fontes de informações disponíveis.

Algumas tecnologias foram desenvolvidas ao longo dos últimos anos para a Web Semântica, e entre as funcionalidades criadas está a utilização de metadados para descrever e localizar a informação na Internet. Comunidades especializadas em diversas áreas do conhecimento têm desenvolvido diversas iniciativas com o intuito de criar esquemas de metadados para a padronização da descrição de recursos utilizados nas diversas áreas do conhecimento, como governo, educação, linguística, processos industriais, agricultura, música, etc.

Definir padrões, entretanto, transcende as questões que têm sido estudadas e apontadas nas áreas da Web Semântica e da Ciência da Informação. Como tratar da complexidade constituída pelo entrelaçamento das questões culturais, sociais e históricas que vão além dos estudos desenvolvidos e apontam para um conseqüente sistema de alta complexidade torna-se um problema com diferentes facetas e constitui-se em cenário em que, no artigo, procurou-se situar e responder aos seguintes questionamentos: (1) como relacionar uma estrutura de ligação que consiga apresentar e relacionar os dados, construída com um conjunto de termos de ontologias distintas?; e (2) como padronizar e estruturar os *Resource Description Framework* (RDFS)?

A primeira questão referente ao relacionamento das estruturas de ligação, para representar e relacionar os dados, depende do conjunto de ligações na estrutura de um RDF, que tem características facilitadoras da fusão dos dados e utiliza as URIs para nomear a relação entre os recursos e a utilização desse modelo. Esse conjunto de ligações permite que os dados estruturados ou semi

estruturados possam ser mixados, expostos e compartilhados entre diferentes aplicações.

Em relação à segunda questão, entende-se que padrões são usados em muitas áreas como ‘modelos’ ou descrições abstratas que reúnem as melhores práticas de algum campo. Verificou-se que a introdução de padrões em Ciência da Computação tem sido, geralmente, atribuída à inspiração tirada do campo da arquitetura, apresentada por Christopher Alexander, em 1987 (GANGEMI; PRESUTTI, 2004). Padrões de *software* tornaram-se mais conhecidos como *Design Patterns*, referindo-se aos padrões comumente conhecidos como GoF (*Gang of Four*), publicado em 1995 por Erich Gamma, Richard Helm, Ralph Johnson e John Vlissides, no intitulado *Design Patterns: Elements of Reusable Object-Oriented Software*. No trabalho, são apresentados 23 padrões para o desenvolvimento de projetos de *software*, muito utilizados em engenharia de *software* (GAMMA et al., 1995).

O termo guarda-chuva *design* e suas derivações, aplicadas a este universo dos padrões, produzem clareza, precisão e eficiência na comunicação de ideias complexas. O *design* é um elemento criador de sentidos, que estrutura e expressa como um ponto de intersecção, por meio de sínteses entre codificações e disciplinas de informação e de comunicação, de linguagens, de estéticas, de disciplinas ligadas à percepção, à cognição e ao comportamento, aos negócios, leis e tecnologias de produção dando suporte aos objetivos dos agentes e criadores.

Outra característica notável de design da informação é o grande número de domínios que se associam com ele. Há grupos significativos, por exemplo, entre os designers gráficos, editores e escritores, e arquitetos de informação onde todos os domínios fazem alguma reivindicação para o campo do design da informação. Tipicamente, as disciplinas são fáceis de definir em pelo menos uma forma de base tática. Os designers gráficos fornecem soluções visuais. Editores e escritores fornecem soluções escritas. Arquitetos de informações fornecem soluções de design estrutural. O design da informação ostensivamente se resume a um amplo conjunto de resultados de informação, não qualquer tipo único ou componente específico de outras disciplinas. O design da informação é um campo integrador (KNEMEYER, 2003, tradução nossa).

Podendo ser pensados como produtos de *design*, embora alguns padrões e estruturas de RDF e de ontologias tenham sido investigados e desenvolvidos, poucos trabalhos concentram esforços para determinar sua utilização e eficácia.

A comunidade de pesquisadores da Web Semântica desenvolveu os padrões ODPs para amenizar algumas complexidades, no que concerne o desenvolvimento de ontologias e criação dos RDFs, favorecendo o reuso e as melhores práticas das codificações utilizadas na área. O entendimento do RDF é essencial para a compreensão da linguagem *Web Ontology Language* (OWL), e essas duas estruturas formam a base do *framework* conhecido como *Ontology Design Pattern* (ODP). Dessa maneira, apresenta-se, de maneira resumida, nas subseções que seguem a linguagem OWL e a estrutura básica RDF.

2 Web Ontology Language (OWL) e Resource Description Framework (RDF)

A linguagem OWL é uma linguagem projetada para representar o conhecimento, rico e complexo, seja este relacionado à ciência ou até mesmo as questões do cotidiano. A linguagem é baseada em lógica computacional e pode ser explorada por programas de computador (WORLD WIDE WEB CONSORTIUM, 2015). Os documentos OWL são baseados em ontologias que, tradicionalmente, têm sido definidos como o estudo de diferentes tipos de entidades e as relações que as entidades suportam entre elas (SMITH, 2004).

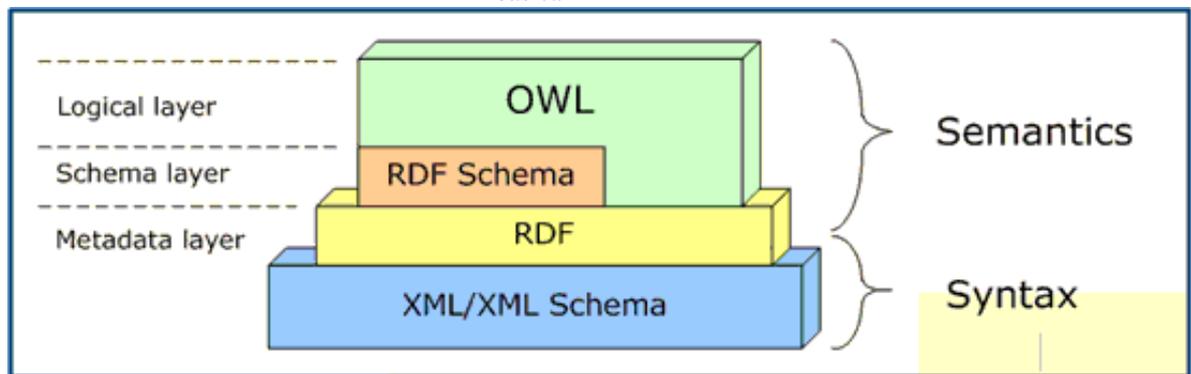
2.1 Web Ontology Language (OWL)

Os documentos OWL são utilizados por meio de algumas tecnologias e, apesar de definir algumas propriedades, são normalmente referidos e/ou encaminhados por algumas linguagens definidas pelo World Wide Web Consortium (W3C), que inclui padrões como: XML, RDF, RDFS, SPARQL, entre outros. A seguir apresentamos a Figura 1, em que podemos perceber os níveis de estrutura básica da linguagem OWL.

A linguagem padrão OWL encontra-se em sua segunda versão, conhecida como OWL2, publicada em 2009, com atualização realizada em 2012 (WORLD WIDE WEB CONSORTIUM, 2012). É possível encontrar um roteiro com as principais propriedades utilizadas para o desenvolvimento de um

documento OWL2, no endereço eletrônico <http://www.w3.org/2002/07/owl> (WORLD WIDE WEB CONSORTIUM, 2012; 2015), que descreve algumas classes e propriedades formadoras da base sintática de um arquivo RDF/XML. Como o roteiro visa somente a criação das necessárias bases sintáticas e semânticas, os termos apresentados nessa base devem ser aprofundados de acordo com outras normas e especificações completas, de acordo com as ontologias desenvolvidas para um dado campo do conhecimento no qual está sendo criado o esquema.

Figura 1 – OWL na Arquitetura semântica da Web/ Níveis superiores de estrutura básica



Fonte: Djurić et al. (2005, p. 211).

Neste estudo, considera-se ontologia como uma moldura que visa normalizar as terminologias utilizadas para a estruturação dos recursos e agentes. Além da possibilidade de variar dentro das generalidades de certo termo, tornando possível a interoperabilidade e o acesso aos repositórios de dados (*Datasets*) de diferentes tipos que estão sendo apontados e acessados pelas linguagens e pelos recursos utilizados em dado projeto. Santarem Segundo e Coneglian (2015, p. 227) indicam que “Para o uso como tecnologia da Web Semântica, entende-se as ontologias como: artefatos computacionais que descrevem um domínio do conhecimento de forma estruturada, através de: classes, propriedades, relações, restrições, axiomas e instâncias.”

Considera-se, também, que se a principal função da utilização de ontologias ocorre para que a estrutura interna das codificações possa se articular entre os repositórios, para que seja possível a existência da interoperabilidade entre os diferentes repositórios que têm informações distintas, esses devem

cruzar as informações e os recursos. Um dos principais problemas da adoção de um conjunto (por vezes muito geral) de categorias básicas de objetos e/ou propriedades e classes orientadas ao objeto é determinar as relações que estas mantêm entre as diferentes categorias na ontologia utilizada.

É importante, portanto, que cada equipe de desenvolvimento de um esquema OWL forneça as informações de forma eficiente e realize uma pesquisa diretamente com os especialistas dos mais diversos campos e sub-campos de conhecimento em que está inserido o desenvolvimento do projeto. E ainda, ter a certeza de que tais informações encontram-se nos repositórios apontados e possam estar acessíveis em tempo real.

Além disso, como apontado por Berners-Lee (2001), em seu exemplo clássico de busca de informação, as estruturas de *design* criadas na Web Semântica permitem que os agentes computacionais sejam capazes de operar de forma eficiente. Ademais, permitem que apresentem as informações e suas relações não somente sintáticas, mas também a níveis profundos e complexos, atingindo as questões pragmáticas e semânticas e apresentando, assim, as informações por meio de um conjunto de regras de inferências.

Desta forma, os agentes computacionais serão capazes de administrar um raciocínio automatizado para a representação do conhecimento e relacionar os dados fornecidos entre os diversos recursos e objetos digitais utilizados. E permite, assim, uma comunicação efetiva e dinâmica entre os agentes e os diferentes sistemas.

2.2 Resource Descriptor Framework (RDF)

O RDF é um modelo padrão para o intercâmbio de dados na Web. Apresenta características que facilitam a fusão de dados, além de suportar a evolução dos esquemas sem a necessidade de que todos os dados sejam alterados, sendo assim, é possível realizar apontamentos para a utilização e para a relação de diferentes recursos (WORLD WIDE WEB CONSORTIUM, 2015). O RDF estende-se a estrutura de ligação da Web que utiliza a marcação dos URIs. Ele é um modelo de dados que referencia objetos e como eles estão relacionados; são representados utilizando a sintaxe do XML, o que permite a interoperabilidade.

A estrutura de um recurso RDF é composta por uma tripla (S, P, O). O S refere-se ao sujeito, P entende-se por um predicado e O é o objeto, sendo cada um deles designado por um recurso.

O RDF Schema (RDFS), por sua vez, é o vocabulário que descreve os recursos de forma hierárquica, utilizando classes e propriedades, ou seja, ele estende o vocabulário RDF. É no RDFS que são definidas as propriedades¹ semânticas e as características de dado objeto.

O esquema exhibe algumas classes, propriedades e restrições do núcleo de um RDFS. O W3C tem concentrado esforços para padronizar tudo o que concerne às especificações da Web e, no presente momento, relativo à Web Semântica. Contudo, existe um longo caminho a percorrer para que tais esforços sejam consolidados e as soluções sejam colocadas em prática.

O RDF pode ser definido a partir de três regras simples:

- a) um fato é expresso como um tripleto na forma de assunto, predicado, objeto. Como em uma frase em linguagem natural;
- b) assunto, predicado e objetos são nomes de entidades quer concretas, quer abstratas no mundo real. Nomes são 1) globais e referem-se à mesma entidade em qualquer documento RDF em que apareçam ou 2) locais e entidade a que se referem não podem ser referidos diretamente fora do documento RDF;
- c) os objetos podem ser valores textuais, designados por valores literais.

Uma das questões que apontamos no estudo é o desenvolvimento de recursos e aplicativos que facilitem a implementação de banco de dados RDF nos *websites* e nos repositórios. A seguir apresenta-se um padrão de *design* baseado em ontologias que visa facilitar a implementação, a partir do desenvolvimento e da documentação de ontologias para as mais diversas áreas do conhecimento.

3 *Ontology Design Patterns* (ODP), convergência e complexidade para criação de padrões de ontologias

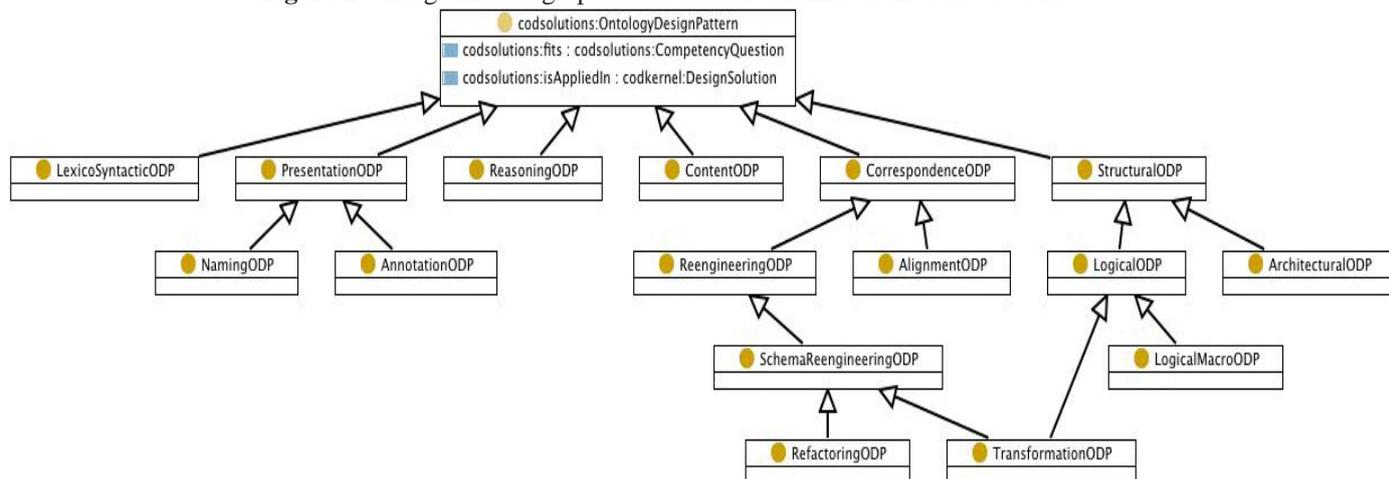
As linguagens utilizadas para a Web Semântica ainda são pouco conhecidas pela comunidade em geral. Isso se dá devido à complexidade para o tratamento de

regras, de recursos, dos agentes, dos sistemas e das possibilidades que esses proporcionam.

O *OntologyDesignPatterns.org* é um portal dedicado ao desenvolvimento e criação de padrões para a Web Semântica, desenvolvidos no contexto do projeto NEON (NEON PROJECT, 2016). O *Ontology Design Pattern* (ODP) é um padrão de *design* baseado em ontologias que, além de apresentar um catálogo público de pesquisa, fornece a semântica e anotações sobre os recursos para facilitar o compartilhamento e futuras atualizações, sendo o catálogo gerado a partir de arquivos OWL (ONTOLOGY..., 2015). Em todo o site, pode ser realizado o *download* de diversos arquivos, o usuário pode escolher a partir de uma lista de padrões, tipos de padrões, domínios, além de disponibilizar informações sobre eventos na área, treinamentos e revisões realizadas pela equipe responsável por meio da *wiki*, na qual cada arquivo OWL disponibilizado descreve um ODP.

Existem vários tipos de ODPs. A equipe de desenvolvimento do ODP identificou os diversos tipos de ODPs, que foram agrupadas em seis famílias, conforme apresentado na Figura 2.

Figura 2 – Diagrama de agrupamento das seis famílias de modelos de ODPs



Fonte: Ontology Design Patterns types (2015).

Descrevemos os padrões apresentados da seguinte maneira:

- a) **ODPs Estruturais** - são divididos em dois sub-padrões: *Ontology Pattern Lógicos* são expressos apenas em termos de um vocabulário lógico, são padrões independentes de um domínio de interesse específico

que podem ser de dois tipos: macros lógicos e padrões de transformação; OPs **Arquitetônicas** afetam de forma geral a ontologia: o seu objetivo é restringir 'como a ontologia deve ser semelhante' de outras propostas. ODPs arquitetônicos surgiram como opções de *design* motivados por necessidades específicas como, por exemplo, restrições de complexidade computacional. Os padrões de ontologias arquitetônicos são úteis como documentação de referência para aqueles estão iniciando o projeto de uma ontologia;

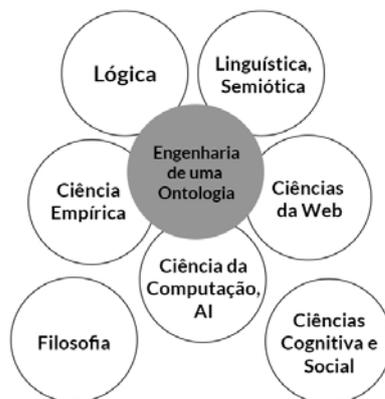
- b) **ODPs de Correspondência** - incluem OP de **Reengenharia** e OP de **Alinhamento**. OP de reengenharia fornece aos designers soluções para o problema de transformar um modelo conceitual, que pode até mesmo ser um recurso não ontológico, para a criação de uma nova ontologia. OP de **Alinhamento** são padrões para a criação de associações semânticas entre duas ontologias existentes;
- c) **ODPs de Conteúdo** - são Padrões de Conteúdo (PC) que codificam os conceitos, em vez de apresentar padrões de projeto lógicos. Em outras palavras, enquanto OP lógicos resolvem problemas de *design* de forma independente de uma conceitualização particular, PCs propõem padrões para resolver problemas de *design* para as classes de domínio e propriedades que permeiam uma ontologia. Portanto, buscam resolver os problemas de conteúdos. Dessa forma, eles têm um vocabulário não lógico, explícito para um domínio específico de interesse (eles são dependentes do conteúdo). PCs fornecem soluções para problemas de modelagem de domínio e afetam somente a região específica da ontologia que lida com tais problemas de modelagem de domínio;
- d) **ODPs de Raciocínio** - Os Padrões de Raciocínio são aplicações de OP lógicos orientados para obter determinados resultados de raciocínio, com base no comportamento implementado em uma máquina de raciocínio. Exemplos de OP de Raciocínio incluem: classificação, subsunção, herança, materialização, e anonimato, etc. OP de raciocínio, quando declarados no topo de uma ontologia, informam sobre o estado em que está a ontologia e permitem que um sistema decida o raciocínio que

precisa ser executado sobre a ontologia, a fim de realizar consultas, avaliação, etc. Exemplos de OP de Raciocínio são as normalizações;

- e) **ODPs de Apresentação** - lidam diretamente com a usabilidade e legibilidade de ontologias a partir da perspectiva do usuário. Elas são propostas como boas práticas que suportam a reutilização de padrões, facilitando a sua avaliação e seleção. Apresentam-se em duas subcategorias: ODPs de nomeação e ODPs de anotação;
- f) **ODPs léxico-sintáticas** - são estruturas linguísticas ou esquemas que consistem em certos tipos de palavras com uma sequência e uma ordem específicas e que permitem generalizar e extrair algumas conclusões sobre o significado que elas expressam. São úteis para associar OP lógicos e de conteúdo simples, como, por exemplo, frases em linguagem natural para fins didáticos.

Para publicar um novo padrão de ontologia diretamente no *website* do *Ontology Design Pattern* é necessário atender alguns requisitos para posterior homologação e inserção no *dataset* do *LinkedData*, por meio de contato direto com o *DataHub*. De acordo com Gangemi et al. (2007), criador dos requisitos e pesquisador do projeto NEON para a criação de um projeto baseado no *design* de ontologias (padrão ODP apresentado), é preciso considerar também o contexto cultural das ontologias, conforme apresentado na Figura 3.

Figura 3 – Contexto cultural para construção de uma ontologia



Fonte: Adaptado de Gangemi et al. (2007, tradução nossa).

Um padrão pode ser de diferentes tipos e inclui:

- a) **lógica** - normalmente, fornece soluções para resolver os problemas de expressividade, por exemplo, expressando relações e cenários entre repositórios realizando a interoperabilidade entre os arquivos OWL;
- b) **arquitetura** - descreve a forma geral de uma ontologia (interna ou externa);
- c) **conteúdo** - são pequenas ontologias que tratam de um problema de modelagem específica, e pode ser importado de uma outra ontologia já desenvolvida;
- d) **apresentação** - oferecem boas práticas para, por exemplo as convenções de nomeação; entre outras questões relevantes (POVEDA; SUÁREZ-FIGUEROA; GÓMEZ-PÉREZ, 2009).

Segundo Gangemi et al. (2007), um bom projeto de uma ontologia deve responder à questões como: sobre o que estamos falando?; por que queremos falar sobre esse assunto?; onde podemos encontrar referências para este tipo de conhecimento? Ao responder tais questionamentos, constitui-se o espaço em que vamos tratar do problema. Nesse sentido, os desenvolvedores e *designers* precisam encontrar o espaço para localizar, referenciar, interoperar e solucionar esses problemas.

Os desafios técnicos e tecnológicos propostos pelos sistemas e pelas redes convergentes – tecnologia que converge em uma só rede de compartilhamento voz, dados, imagens, sensores, etc. – necessitam ser enfrentados para que estas sejam efetivas como construtoras de conhecimento e para que possam enfrentar desafios de caráter sócio-técnicos. Por outro lado, é necessário escalabilidade para que as redes convergentes sejam efetivas e alcancem a um grande número de internautas.

Além disso, um sistema escalável necessita manter seu desempenho, independentemente de suas extensões ou da utilização dos seus recursos, e estar em beta perpétuo, permitindo frequentes atualizações sem desligamento, para suportar e aumentar processamentos que se fizerem necessários perante a adição

de processadores e de dispositivos de armazenagem. E ainda a escalabilidade horizontal, possibilitando a adição de mais nós aos seus subsistemas.

A noção de um mundo construído por tais redes conectivas convergentes impõe o questionamento sobre a pluralidade de formas de percepção fenomenológica, a partir do descobrimento das diversas manifestações da matéria – tanto no nível mecânico newtoniano, quanto no nível quântico. E é válido para escalas menores, regidas por leis não cartesianas-newtonianas.

No entanto, propõe também uma questão que permanece, ainda hoje, sem resposta definitiva: sobre a capacidade de pensamento das máquinas. Em *Computing Machinery and Intelligence*, publicado pela primeira vez em 1950, na *Mind*, Alan Turing criou um desafio de predição em que um sujeito à distancia, em um terminal de computador, deveria decidir se o seu interlocutor em outra ponta do experimento era um humano ou uma máquina. Conhecido como o Teste de Turing, a experiência discutia a Inteligência Artificial (IA), baseada na competência linguística maquínica, utilizando-se do método cartesiano. O problema apontado nesse teste, entretanto, refere-se à perspectiva essencialmente humana do pensamento e inteligência. Das discussões, surgiram duas tendências acadêmico-científicas quanto à inteligência artificial (IA).

Uma dessas tendências preconizava uma IA forte, em que se considerava uma forma de inteligência computacional capaz de raciocinar e de resolver problemas de forma autoconsciente. Consciência, entretanto é diferente de inteligência; a consciência não deriva de sistemas lógicos, baseados em regras, mas uma das condições apontadas para a emergência da consciência é fazer parte de um corpo físico com propriocepção (JORENTE, 2014, p. 120).

Contudo, deve se considerar que nas redes conectivas convergentes, também, o entrelaçamento entre objeto, observador e método em um sistema relacional de complexidades que predispõe à emergência por meio das estruturas de linguagens e de codificações. São as conexões e nós da rede que fazem interagir os padrões de maneira complexa, lembrando que complexidade aqui significa a soma das partes e a emergência resultante entre tal soma. Levando em conta essa perspectiva, a outra tendência propunha uma tentativa de replicação da inteligência humana que caracterizou a denominada IA fraca.

Na IA fraca, meios de processamento da linguagem natural reproduziam formas de funcionamento da linguagem humana (excluindo-se as questões da consciência), para criar programas computacionais como o *Artificial Linguistic Internet Computer Entity* - A.L.I.C.E. A ideia de uma IA fraca migrou para a compreensão da complexidade sistêmica no A.L.I.C.E. (JORENTE, 2014, p. 120).

Simulando formas de conversação humana, o A.L.I.C.E usava de programação em Java e resultou na *Artificial Intelligence Markup Language* (AIML), linguagem baseada em *eXtensible Markup Language* (XML), específica para tais programas e extensiva aos clones Alicebots. O desenvolvimento da AIML, entre os anos de 1995 e 2002, possibilitou criar diálogos semelhantes à linguagem natural, por meio de *softwares* que procuravam simular também o funcionamento da inteligência humana, tendo em vista a complexidade da linguagem.

A interação entre os sujeitos da comunicação humanos e maquímicos por meio dos artifícios linguísticos é, portanto, um dos aspectos mais inovadores da linguagem AIML, cuja ideologia precedeu o que hoje é realidade em programas interativos e, também, na criação de padrões de *design* para ontologias. Assim, o princípio que rege a ideia e o *Design* de Ontologias é o da emergência, advindo conceitualmente da Teoria da Complexidade.

A emergência é responsável pela formação de padrões de complexidade variada resultantes de interações simples. Se nos processos cerebrais observados, neurônios sem consciência individual em si mesmos podem, quando em interação, fazer emergir pensamento, consciência e comportamentos complexos, um sistema, humano ou não, necessita um nível de diversidade, organização e conectividade para provocar comportamentos emergentes. Por outro lado, se a informação necessita ser comunicada para reelaborar-se em conhecimento, o seu apoio ou suporte representacional necessita ser apreendido como processo tecnológico que é - escrito, imagético, sonoro ou multimidiático, já que os suportes de informação determinam em muito as suas formas de acumulação ou de estocagem e, também, de abordagem para sua compreensão.

O universo da informação, em seus aspectos materiais e imateriais que propõem formas concretas de leitura, assim como o dos sujeitos que com ela

interagem, envolvem-se como partes do sistema determinante do conhecimento possível pelas possibilidades e limites que impõem às interações. Esses universos devem convergir para o equilíbrio na real apropriação dos materiais informativos.

Os processos digitais informacionais contemporâneos - que compreendem novas formas de relacionamentos entre diferentes subsistemas de informação - criam e ampliam de maneira transdisciplinar novas possibilidades de reflexão e de exploração do conhecimento neles veiculado. (JORENTE, 2011, p. 11).

Portanto, nos ambientes digitais conjugam-se protocolos tecnológicos, capacidades cognitivas diversas, competências oportunizadas e convenções sociais, entre outros subsistemas da complexidade que constituem a comunicação digital atualmente. Os protocolos tecnológicos referem-se às interações possibilitadas pelos aparatos e *softwares*. As diversidades das capacidades de cognição devem ser consideradas e potencializadas juntamente com as competências por equipes multidisciplinares, envolvidas na preparação da informação para as interações, bem como as convenções sociais, para que, com a ampliação dos hábitos de *linkagem* (processo dinâmico hipertextual de navegação) e a crescente disponibilização de acesso e de engajamento nos ambientes digitais, de forma idealizada, possa derivar benefícios sociais.

[...] estamos somente arranhando a superfície do que poderia ser alcançado com uma investigação científica mais profunda de seu design, operacionalidade e impacto sobre a sociedade. A Web permanece uma plataforma universal: independentemente de qualquer instrumento de hardware específico de uma plataforma de software, de linguagem, cultura ou falta de habilidades (BERNERS-LEE, 2007, tradução nossa).

Ambientes digitais de aprendizado e entretenimento conclamam a cooperação e o empoderamento quando propõem elementos da Web colaborativa. Entre tais elementos estão a participação e a conversação, o auto-arquivamento, os códigos visualizáveis e a contextualização, permitindo o beta perpétuo das constantes melhorias sistemáticas. Isso se deve à característica de configuração rizomática da rede mundial que se configura como sistema complexo, aberto, dinâmico, fluído e compreensivo. Com tais características a

Web colaborativa propõe-se como um ambiente complexo para a expressão da diversidade cultural, para a prestação de serviços on-line e, principalmente, para o desenvolvimento de práticas educacionais inclusivas que necessitam ser implementadas para o salto quântico que o sistema permite.

4 Considerações

No novo paradigma apresentado como pós-custodial, os profissionais da informação necessitam mudar o foco de atenção, pensando no redesenho dos sistemas digitais por meio do *design* da informação.

Em termos gerais de documentação, os *design patterns* são altamente estruturados. A exemplo das estruturas, foi apresentado o diagrama de agrupamento com as seis famílias de ODPs propostas no projeto NEON. No entanto, em linhas gerais, são documentados a partir de um modelo que identifica a informação necessária para entender o problema do *software* e/ou da ontologia e a solução em termos de relacionamentos entre as classes e objetos necessários para implementar essa solução. Não há um ponto comum na comunidade de *design patterns* sobre como descrever um *template* de *patterns*. Alguns autores preferem ser mais expressivos e menos estruturados, enquanto outros preferem que seus *templates* sejam mais precisos e altamente estruturados.

Conclui-se que a rede mundial de computadores assiste ao nascimento de uma nova formação de inteligência, a *noomorfosis* digital (do grego *noos*=inteligência, *morphosis*=formação), identificada como um fenômeno transcendental ainda imperceptível como um todo, mas já aparente em diversas situações que envolvem um número crescente de nativos digitais. Tais nativos já nasceram imersos sob as novas formas de criação e disseminação da informação da contemporaneidade, com uma inteligência funcionalmente distinta, aparelhada com capacidades necessárias, especialmente desenvolvidas para viver e operar no Novo Entorno Tecnossocial (NET), gerado por esta tecnologia (SÁEZ VACAS, 2006).

A natureza conceitual da rede mundial de computadores é a de convergência com os sujeitos que nela navegam via aparatos tecnológicos que

permeiam a percepção da informação estruturada para registrar e organizar o mundo e para, também, proporcionar o acesso ao conhecimento. A presença da tecnologia no cotidiano das pessoas torna a investigação dessa temática na Ciência da Informação importante para uma formação reflexiva dos profissionais da informação na gestão dos recursos informacionais alocados, nos mais diversos suportes e ambientes.

Referências

BERNERS-LEE, Tim. The semantic web. **Scientific American**, New York, May 2001.

BERNERS-LEE, Tim. Testimony of sir Timothy Berners-Lee. In: DIGITAL FUTURE of the United States: the future of the world wide web – part I. [S.l.]: Committee on Energy and Commerce, 2007. Disponível em: <<http://dig.csail.mit.edu/2007/03/01-ushouse-future-of-the-web>>. Acesso em: 15 ago. 2015.

DJURIC, Dragan et al. **MDA-based ontological engineering**. World Scientific Publishing: Singapore, 2005. p. 203-231. Disponível em: <<http://devedzic.fon.rs/publications/HANDBOOK2005.pdf>> Acesso em: 18 jul. 2017.

GAMMA, Erich et al. **Design patterns: elements of reusable object-oriented software**. [S.l.]: Addison-Wesley, 1995.

GANGEMI, Aldo; PRESUTTI, Valentina. Ontology design patterns. In: STAAB, Steffen; STUDER, Rudi (Ed.). **Handbook on ontologies**. 2nd ed. Berlin: Springer, 2004. p. 221-244.

GANGEMI, Aldo et al. **Towards a catalog of OWL-based ontology design patterns**. 2007. Disponível em: <http://oa.upm.es/5212/1/Towards_a_Catalog_of_OWLbased_Ontology_Design_Patterns.pdf>. Acesso em: 28 ago. 2015.

HEATH, Tom; BIZER, Christian. **Linked data: evolving the web into a global data space**. San Rafael: Morgan & Claypool, 2011.

JORENTE, Maria José Vicentini. Cultura da Wikipédia como política de capacitação compartilhada em meio digital de socialização do conhecimento. **Tendências da Pesquisa Brasileira em Ciência da Informação**, Belo Horizonte, v. 4, n. 1, p. 53-90, 2011.

- JORENTE, Maria José Vicentini. Design da informação, linguagens convergentes e complexidade na rede social e ambiente digital do Facebook. **Informação & Tecnologia**, João Pessoa, v. 1, p. 116-129, 2014.
- KNEMEYER, Dirk. **Information design: the understanding discipline**. 2003. Disponível em: <http://online.sfsu.edu/jkv4edu/2DMG/projects/infodesign_boxarrows.pdf>. Acesso em: 5 set. 2015.
- MOLES, Abraham. **Sociodinâmica da cultura**. São Paulo: Perspectiva. 1974.
- NEON PROJECT. [S.l.]: MediaWiki, 2016. Disponível em: <<http://www.neon-project.org/nw>>. Acesso em: 10 ago. 2015.
- ONTOLOGY Design Patterns.org. 2015. Disponível em: <<http://ontologydesignpatterns.org>>. Acesso em: 5 set. 2015.
- POVEDA, María; SUÁREZ-FIGUEROA, Mari Carmen; GÓMEZ-PÉREZ, Asunción. Ontology analysis based on ontology design patterns. In: WORKSHOP ON ONTOLOGY PATTERNS, 3., 2009, Boston. **Proceedings...** [S.l.]: CEUR-WS.org, 2009. Disponível em: <<http://ceur-ws.org/Vol-516/pap05.pdf>>. Acesso em: 15 ago. 2015.
- SÁEZ VACAS, Fernando. **Noomorfosis digital**. 2006. Disponível em: <<http://antoniofumero.blogspot.com.br/2006/08/noomorfosis-digital.html>>. Acesso em: 20 ago. 2015.
- SANTAREM SEGUNDO, José Eduardo; CONEGLIAN, Caio Saraiva. Tecnologias da web semântica aplicadas a organização do conhecimento: padrão SKOS para construção e uso de vocabulários controlados descentralizados. In: GUIMARÃES, José Augusto Chaves; DODEBEI, Vera. (Org.). **Organização do conhecimento e diversidade cultural**. Marília: FUNDEPE, 2015. p. 224-233.
- SCHMACHTENBERG, Max; BIZER, Christian; PAULHEIM, Heiko. **State of the LOD cloud 2014**. 2014. Disponível em: <<http://linkeddatacatalog.dws.informatik.uni-mannheim.de/state/>>. Acesso em: 30 ago. 2016.
- SMITH, Barry. **Beyond concepts: ontology as reality representation**. 2004. Disponível em: <<http://www.ontology.buffalo.edu/ontology>>. Acesso em: 10 jun. 2015.
- WORLD WIDE WEB CONSORTIUM. **OWL 2 Web Ontology Language RDF-Based Semantics (Second Edition)**: W3C Recommendation. 11 December 2012. Disponível em: <<https://www.w3.org/TR/owl2-rdf-based-semantics/>>. Acesso em: 18 jul. 2017.

WORLD WIDE WEB CONSORTIUM. **RDF Schema 1.1**: W3C recommendation 25 february 2014. 2014. Disponível em: <http://www.w3.org/TR/rdf-schema/#ch_summary>. Acesso em: 10 ago. 2015.

WORLD WIDE WEB CONSORTIUM. 2015. Disponível em: <<http://www.w3c.org>>. Acesso em: 10 ago. 2015.

Design of patterns in the Semantic Web: prospects and challenges

Abstract: The article discusses the creation of ontology standards and how the resources and practices developed by the Ontology Design Patterns team (NEON Project) propose standards and recommendations for the organization, access and retrieval of resources in the Semantic Web. Thinking about the complexity of creating ontology-based language patterns involves understanding that this is expressed as a concern, self-referential question regarding its convergent nature. To answer it, one needs to understand that the main quality of what is complex, is combining many elements or parts. From this point of view, there are several standards and models for the creation of 'design patterns', based on content, language, knowledge and logical patterns, among others. This exploratory and descriptive research aimed to elucidate standards and norms of ontology design, by providing criteria and methods on the object, to formulate the described hypothesis, and to identify the factors and variables related to the studied practices. We concluded that the investigation of this issue in Information Science is important for reflexive training of information professionals.

Keywords: Ontology Design Patterns. Semantic Web. Complex Systems. Convergences.

Recebido em: 04/01/2017

Aceito em: 10/03/2017

¹ Uma lista com as classes básicas e propriedades utilizadas nos RDFS pode ser encontrada na página *RDF Schema* da W3C (WORLD WIDE WEB COUNSORT, 2014).