

A relação das enciclopédias com os índices e a *Web* semântica: linhas de força para a organização e significação na pós-modernidade

The relation of encyclopedias with the indexes and the semantic Web: Organization and signification supports in postmodernity

Silvana Drummond MONTEIRO¹

Resumo

O estudo objetiva estudar o processo de significação (ou “semantização”) que ocorre no ciberespaço, especialmente, por meio da Wikipédia, da *Web* semântica e dos bancos de conhecimento – *Knowledge Databases* – como o *Freebase* e o *DBpedia*, para citar os mais importantes, que modelam os dados com *Resource Description Framework*. Por meio de pesquisa documental, de revisão, buscou-se compreender esse movimento que constrói, no ciberespaço, uma rede ou teia de significações para ser recuperada, ativada e interpretada por máquinas e humanos. A partir da análise do *corpus* teórico, percebe-se a concepção de enciclopédia como Educação, depois como repertório de conhecimento, no mundo clássico, e como obra de referência, na modernidade. Na pós-modernidade ganha um novo plano de consistência, como organização semântica do saber e como banco de dados. No ciberespaço, os índices contemporâneos e a Wikipédia são um *corpus* em contínua renovação e expansão e estão interligados por metadados de representação semântica, que povoam as entidades dos bancos de conhecimento legíveis pelos agentes inteligentes. Essas relações colocam em ressignificação conceitos de enciclopédia, índices e até de Semântica, evidenciando os agenciamentos maquínicos e híbridos dos processos de organização e significação na pós-modernidade.

Palavras-chave: Enciclopédias. Índices. *Web* semântica. Wikipédia.

Abstract

The study aims to study the process of signification (or “semantization”) that occurs in cyberspace, especially through Wikipedia, the semantic web and Knowledge Databases as Freebase and DBpedia, to name the most important ones, modeling data with Resource Description Framework. Through documentary research review, we sought to understand this movement that builds, in cyberspace, a network or web of significations to be retrieved, activated and interpreted by machines and humans. From the theoretical corpus analysis, we can see the design of encyclopedia as Education, then as a repertoire of knowledge, in the classical world and as a reference work, in modernity. In postmodernity it gets a new plane of consistency, as semantic organization of knowledge and as a database. In cyberspace, contemporary indexes and Wikipedia are a corpus in continuous renewal and expansion and are interconnected by semantic metadata representation, which fill the corpus of knowledge databases readable by intelligent agents. These relationships reframe concepts of encyclopedia, indexes and even Semantics, showing machinic and hybrids assemblages of the organization and signification processes in postmodernity.

Keywords: *Encyclopaedias. Indexes. Semantic web. Wikipedia.*

¹ Universidade Estadual de Londrina,, Departamento de Ciência e Informação, Programa de Pós-Graduação em Ciência da Informação. Rod. Celso Garcia Cid, PR 445, Km 380, s/n., *Campus* Universitário, 86057-970, Londrina, PR, Brasil. E-mail: <silvanadrumond@gmail.com>.

Recebido em 30/6/2015, reapresentado em 26/2/2016 e aprovado para publicação em 13/9/2016.

Introdução

Pensar na enciclopédia parece ser simples, em um primeiro momento, mas basta estudá-la para verificar a complexidade de seu conceito. Sem que se perceba, dada a associação que dela se faz com o projeto editorial do Iluminismo, que por sua vez fez parte do ideário da modernidade, ela se constitui em entidade polissêmica.

Seria pertinente falar em enciclopédia, produto do Iluminismo, em época pós-moderna marcada pelas relações híbridas (LATOURE, 1994) e os consequentes agenciamentos maquínicos (DELEUZE; GUATTARI, 1995-1997)? A pragmática dos índices contemporâneos gerados pelos mecanismos de busca, a robustez da Wikipédia e a epígrafe de Lyotard (1988) autorizam a pensar que sim. Mas em que condições e em qual contexto são as reflexões que norteiam este artigo.

Desde a sua associação com a Educação (concepção grega) e depois com o repertório de conhecimento (concepção romana), passando pelas relações mais complexas, como representação semântica geral do saber e/ou do mundo (ECO, 2013), até a sua ligação contemporânea (da Wikipédia) (VRANDEËIAE; KRÖTZSCH, 2014) com as tecnologias da Web Semântica: *Resource Description Framework* (RDF); *Ontology Web Language* (OWL); *eXtensible Markup Language* (XML); os *Knowledge Databases* (ou bancos de conhecimento livres como: o *Wikidata*; o *BPedia* (LYOTARD, 1988); o *Knowledge Graph* do *Freebase*, os dois últimos do *Google*), o percurso é longo, embora possível o esboço e o esforço de pensar a organização do conhecimento e a significação no ciberespaço.

São objetos que margeiam a representação e a organização do conhecimento (no sentido *lato sensu*): sejam os mecanismos de busca, seja a enciclopédia, como método ou sistema de organização do conhecimento, ambos com foco no ciberespaço e suas manifestações simbólicas e a produção de sentido.

Para ajudar nessa trajetória, dois autores são essenciais ao *corpus* teórico do artigo: Eco (2013); Pombo (2002; 2003a). Ainda que na próxima seção o tom revisional esteja acentuado, mister à compreensão e ao desenvolvimento do assunto perceber que, para além

da ideia de enciclopédia como produto editorial (fonte de informação, obra de referência etc.), ela abriga construtos que ora escondem, ora revelam visões de um mundo que se quer ordenado, de alguma maneira, em todas as épocas. Em direção oposta, a *Web Semântica* não terá abordagem histórica e tampouco revisional, mas com ênfase em sua manifestação atual por meio de sua relação com os objetos deste texto.

Seguindo Burke (2003), a enciclopédia é uma parte da história social do conhecimento e, nesse sentido, os índices, as bibliotecas, os arquivos e os museus também o são, embora o artigo faça menção apenas aos dois primeiros.

A pesquisa tem delineamento documental, visando relacionar esses temas que estão fortemente entrelaçados no ciberespaço. Diante da grande ambiguidade sobre sua definição, “A pesquisa documental é aquela cujos objetivos ou hipóteses podem ser verificados através da análise de documentos bibliográficos ou não bibliográficos [...]”. Dessa forma, não deve ser confundida com busca de informação ou com levantamento bibliográfico (CASTRO, 1977, p.76).

O artigo em tela consiste em uma primeira etapa de investigação sobre a abordagem semântica dos índices contemporâneos e tem como objetivo apresentar, dentro do espaço disponível e possível, a relação entre a *Web Semântica*, preconizada por Bernes-Lee *et al.* (2001), e sua atualização e aplicação a partir da enciclopédia pós-moderna, no caso a Wikipédia, e os mecanismos de busca, doravante denominados índices contemporâneos.

Algumas definições em contextos históricos: o esforço da organização do conhecimento por meio da enciclopédia

Como é difícil apresentar algumas definições de enciclopédia descoladas de seu contexto histórico e pragmático, esta seção empreenderá comentar essas ligações, de forma reconhecidamente breve, ao longo do projeto enciclopedista: clássico (grego e romano); medieval; moderno; e contemporâneo² (discutido

² Batista e Rodrigues (2014) utilizam o termo “pós-contemporâneo” para expressar o momento atual da sociedade, com as características do pós-moderno.

também neste artigo como pós-moderno³, em razão da extensa história da enciclopédia, que é indissociável do projeto de conhecimento humano que inclui, por sua vez, o ensino, a visão de mundo, o conhecimento, os livros, as bibliotecas e os autores.

Ao lado de sua definição, pode-se dizer que, historicamente, a enciclopédia alcançou função pedagógica na antiguidade, teológica na Idade Média, e epistemológica na Idade Moderna, visto que serviu de método de organização e, em grande medida, de unificação do conhecimento⁴.

Dada a brevidade histórica sobre a enciclopédia, neste artigo, busca-se a definição de História de Foucault (2007, p.300), em que:

[...] não deve ser entendida como coleta de sucessões de fatos; ela é o modo de ser fundamental das empiricidades, aquilo a partir de que elas são afirmadas, postas, dispostas e repartidas no espaço do saber para eventuais conhecimentos e para ciências possíveis.

Eco (2013, p.34) afirma que o conceito de enciclopédia como compilação do conhecimento em um volume (ou volumes) é um construto romano e medieval, uma vez que para os gregos o repertório do saber tinha equivalência à biblioteca e ao museu: “A atitude enciclopédica se desenvolve, sobretudo em ambiente romano, onde se recolhe todo o saber grego, como numa operação de apropriação do patrimônio [...]”.

Os gregos estavam mais preocupados com a Educação, e não necessariamente com a coleta do conhecimento em compêndio, uma vez que o termo “enciclopédia”, no ambiente de seu nascedouro, significava “educação completa”, como defende Eco (2013), ao interpretar “círculo” como “completo” da palavra grega *Enkyklos Paideia* (literalmente “círculo do conhecimento”), pois:

O mundo helênico designava a função que os romanos e os medievais designarão à enciclopédia não a um volume que fala de todas as coisas, mas a uma reunião de todos os volumes existentes, a *biblioteca*, e a recolha de todas as

coisas possíveis, o *museu* (ECO, 2013, p.34, grifos do autor).

Nesse sentido, Pombo (2003b) sintetiza que o enciclopedismo surge na Grécia tardia “[...] perante uma exigência escolar, um procedimento acadêmico que visa conservar e prolongar pela palavra escrita a palavra dita do professor”.

A enciclopédia, como compêndio, surge com a clássica obra “*História Naturalis*” do latino Plínio (Caius Plinius, 23-79 d.C.). Para Pombo (2003b), a obra que serviu de base para o enciclopedismo possuía 20 000 fatos, quase 500 autores consultados (146 latinos e 327 gregos, em sua maioria), 37 volumes – conhecidos – com os seguintes assuntos: Cosmologia, Vulcanologia, Climatologia e Astrologia (livros I e II); Geografia, Demografia, Etnografia (livros III a VI); Antropologia e Fisiologia humana (livro VII); Zoologia (livro VIII a XI); Botânica, Agricultura, Jardinagem, (XII a XIX); Farmacopeia Natural, Medicina e Magia (do volume XX ao XXXII); Mineralogia, Arquitetura e Artes Plásticas (XXXIII a XXXVII).

Eco (2013) afirma que Plínio explica a lógica de sistematização que reúne os assuntos do “mundo”, por meio do índice da obra. Nesse sentido, sobre a compilação de Plínio, explica:

[...] a enciclopédia não pretende registrar o que realmente existe, mas o que as pessoas tradicionalmente consideram que exista – portanto, tudo aquilo que uma pessoa instruída deveria saber, não só para conhecer o mundo, mas também pra compreender os discursos sobre o mundo (ECO, 2013, p.36).

Justamente por o termo grego *encyclopaedia* significar “círculo do aprendizado”, as enciclopédias continham a organização dos currículos (já no mundo medieval e no moderno) ou daquilo que deveria fazer parte do aprendizado de todo homem civilizado, em face da incipiência do ensino no mundo clássico.

Em suma, no mundo clássico (Século VIII a.C - 476 d.C.), a enciclopédia tem duas acepções: para os gregos é um instrumento para o ensino, e para os latinos, uma

³ O Pós-moderno não é um período da História, como as demais eras supracitadas, mas o Pós-moderno, como temporalidade filosófica, contempla o tom das reflexões, deste artigo, por aduzir a problemática (aqui, não esvaziadora) da sociedade – e do conhecimento – com as Tecnologias da Informação e Comunicação, pois “O ambiente pós-moderno significa basicamente isso: entre nós e o mundo estão os meios tecnológicos de comunicação.” (SANTOS, 2006, p.13).

⁴ A profusão de obras, de compilações, de inspirações de eruditos e filósofos que são atribuídas como enciclopédias, pelos estudiosos, fez com que apenas algumas fossem citadas neste artigo.

compilação do saber existente, embora também fosse utilizada para a formação de seus cidadãos. Pombo (2002) resume essas características no Quadro 1.

Na Era medieval (476-1453 d.C.), as intenções e vocações residiam, especialmente, mas não exclusivamente, na necessidade da interpretação escritural. O enciclopedismo medieval utiliza, como método de organização do conhecimento, o currículo da universidade medieval, ou seja, o *trivium* e o *quadrivium* (BURKE, 2003), e é fortemente influenciado pela Teologia, como característica principal dessa época que tem Deus como centro do mundo.

O método que antecede os sistemas de organização do conhecimento surgidos na Idade Média e na Moderna é a metáfora da árvore (Século XV) como estrutura organizativa do mundo real. Esse instrumento naturalizava as classificações e muitos são os exemplos: a árvore Lógica de Porfírio (Século III d.C., embora apenas sugerisse a imagem da árvore), a árvore do conhecimento de Lúlio (1300), a árvore judiciária de Ludwig Gilhausen (1612), árvores genealógicas ou consanguinidade, da gramática etc. (BURKE, 2003).

Entretanto, no século XVII, outro termo mais abstrato começa a entrar em cena para designar a organização do conhecimento: "sistema", sendo três os principais subsistemas: o currículo, a enciclopédia e as bibliotecas. Em muitos casos, esses subsistemas sobrepunham-se em seus arranjos, pois faziam parte do tripé intelectual da época:

Não devemos supor que qualquer dos três sistemas fosse reflexo não problemático de

categorias mentais ou idéias gerais sobre a organização do conhecimento. [...] Enciclopédias eram produtos vendidos no mercado aberto e sujeitas a suas pressões [...] (BURKE, 2003, p.83).

O neologismo latino "enciclopédia" surge no século XVI, ou seja, no início da Era Moderna (1453-1789 d.C.). De fato, só a partir do século XVII, quando o enciclopedismo se constitui como um verdadeiro movimento, é que o termo se populariza, aproximando-se do seu significado atual (POMBO, 2002):

Beneficiando do retorno humanista aos textos originais, o que vai permitir a eliminação de erros que o enciclopedismo medieval introduzia e perpetuava, estas obras, agora impressas, puderam contar com mais fáceis e rápidas condições de realização e com um público mais alargado e diversificado (POMBO, 2003b).

A diferença das primeiras manifestações de enciclopédias que começaram a ser sistematizadas a partir da estruturação dos currículos, de acordo com Eco (2013), é a consolidação das ciências. Informações sobre magias e lendas foram substituídas pelo conhecimento científico. Acrescenta-se a isso a ordem alfabética, o uso de línguas nacionais em vez do latim e a função constitutiva do enciclopedista (POMBO, 2003b)

O século XVII é considerado o século de ouro das enciclopédias e dentre elas a mais famosa é a "*A Encyclopédie ou Dictionnaire Raisonné des Sciences, des Arts et des Métiers, par une Société de Gens de Lettres*" de Diderot e D'Alembert (1751-1765), doravante *Encyclopédie*.

Apesar de ser arranjada em ordem alfabética e autoria coletiva, sua lógica ainda é disciplinar. Ela é

Quadro 1. Comparação entre o enciclopedismo grego e latino.

Como principais traços comuns ao enciclopedismo antigo poderíamos assinalar o fato de se tratar, em todos os casos

- De compilações de autoria individual;
- Que se constituem como percurso de formação educativa;
- Que, por esse fato, se dirigem a um público relativamente restrito e homogêneo, já detentor de um elevado nível de educação.

Sob a forma de compêndio para fins escolares, a enciclopédia grega recorta-se, como vimos, sobre a incipiente estrutura curricular do não menos incipiente processo de institucionalização escolar.

Mais atento às questões práticas, enfatizando, sobretudo os temas geográficos, de medicina, história e governação, o enciclopedismo latino está ordenado à superior formação dos cidadãos do império apresentando-se como a súpula de tudo aquilo que um homem educado deveria saber ou poderia necessitar.

Fonte: Pombo (2002).

introduzida com um “*Tableau des Connaissances*”, ou seja, uma tabela formando a árvore do conhecimento enciclopédico, em que D’Alembert apresenta o esquema de classificação do conhecimento baseado nas categorias filosóficas (das faculdades humanas) de Bacon: memória, razão e imaginação.

Resumo admirável da ideologia da tolerância que está subjacente à ‘Encyclopédie’! Por outras palavras, a Encyclopédie propõe um discurso contínuo que subjaz à descontinuidade alfabética. Por detrás (ou por baixo) do texto fragmentário da Encyclopédie, corre o rio poderoso e violento de uma textualidade plena e persuasiva (POMBO, 2003b, *online*).

No entanto não se pode ignorar que a ordem alfabética provoca mudanças cognitivas significativas e também reforça a ideia que se tem hoje de enciclopédia, isto é, as obras não são para ser lidas e sim consultadas e, nesse sentido, outra característica haveria de marcar o conhecimento: a ordem alfabética. Ou seja, elas vão motivar (e exprimir) a moderna fragmentação do conhecimento (BURKE, 2003).

Na aceção de obra de referência, a *Encyclopédie* foi concebida em um momento em que ocorriam várias reformas sociais e do conhecimento, de forma que ela também foi uma nova maneira de organizar o conhecimento (BURKE, 2012).

Apesar da compreensão de sua origem, conceito e funcionalidade, Eco (2013, p. 13) apresenta tanto a enciclopédia quanto o dicionário como construções semióticas extremamente complexas. Para o autor, existem dois modelos ou concepções de representação semântica que remetem a uma representação geral do saber ou do mundo: o dicionário e a enciclopédia, porém “O dicionário e a enciclopédia semióticos não são diretamente equiparáveis aos seus congêneres editoriais, mas ao conhecimento de uma língua e ao conhecimento de mundo, respectivamente”.

Em linha teórica, essa discussão a respeito dos signos, significados e conhecimento inicia-se com a lógica aristotélica sobre as definições e o uso generalizado da árvore porfiriana como método ou instrumento lógico para obtê-las.

Confronta, também, a real intenção da definição de Aristóteles *versus* a classificação dos gêneros e es-

pécies por Porfírio: “Por isso uma classificação à maneira de dicionário não serve para definir um termo, mas somente para permitir usá-lo de modo logicamente correto” – por exemplo: cão, gênero *canis*, família *canídeos*, subordem dos *Fissípedes*, da ordem dos carnívoros, da subclasse dos placentários, da classe dos mamíferos (ECO, 2013, p.25). Enquanto o dicionário (que ele relaciona à classificação) é uma compilação de significados, a enciclopédia é uma compilação de saberes, um maquinismo retórico, um sistema semiótico.

Pombo (2002, *online*) também compara o dicionário à enciclopédia e completa:

Ao contrário do dicionário, a enciclopédia tende ao aproveitamento semântico dos recursos diagramáticos da linguagem pondo-os ao serviço da referência e da descrição – imagética, icônica, cartográfica – do mundo para que ela remete e que nela se espelha.

Dessa forma, uma enciclopédia é uma rede semântica de informações “ativáveis” passíveis de descrição, latência e “O formato da rede a ser ativada é prescrito pelos contextos e pelas circunstâncias de enunciação [...]” (ECO, 2013, p.66).

Foram muitas experiências ao longo da civilização ocidental até chegar à Wikipédia: os formatos das enciclopédias; a tensão entre a seletividade e exaustividade; o enfoque disciplinar dominante desde sua primeira manifestação e depois as temáticas e os índices – os quais Eco (2013) designa como metalinguagem e que em determinados exemplos são até mais interessantes que a própria enciclopédia, pois revelam a lógica do arranjo da obra; posteriormente o construto alfabético; a enciclopédia como produto editorial; a enciclopédia tornando-se um índice; e hoje os índices transformando-se em enciclopédias, no ciberespaço.

Assim, Pombo (2003a, p.28) considera que, sendo uma obra “Aberta e aparentemente labiríntica, no recôndito das suas mais escondidas pregas, a enciclopédia continua paradoxalmente a ser um projeto cartográfico”. É justamente essa discussão que interessa para o ciberespaço, seu aspecto complexo linguístico, projeto labiríntico digital, para além da ideia de enciclopédia como produto editorial.

Retornando à visão de Eco (2013) a respeito da enciclopédia sob o enfoque semântico, os signos e as coisas, sua abordagem semiótica e pragmática a torna contemporânea e extremamente pertinente para estudar as manifestações de organização e significação no ciberespaço, conforme apresentado na próxima seção. A reflexão de Eco (2013) sobre a enciclopédia retoma uma discussão cara à Semântica e à Filosofia, ou seja, a ordem das coisas e das palavras. Essa relação, apontada por Pombo (2002, *online*), implica que:

A língua filosófica universal pode e deve ser construída em paralelo com a enciclopédia. Se, por um lado, a construção da língua universal supõe a análise dos conceitos fundamentais e a elaboração das definições primeiras, ou seja, toda a enciclopédia; inversamente, a transposição simbólica dos conteúdos cognitivos e a sua inserção na rede dos signos já constituídos, permite estabelecer novas virtualidades de sentido e, portanto, facilita o avanço da enciclopédia.

O ciberespaço e as enciclopédias digitais: ressignificações na pós-modernidade

Sabe-se que a modernidade trouxe, em seu ideário, os princípios do Iluminismo, que por si só não se contrapõem, em certa medida no caso do antropocentrismo, com a atualidade, tais como o avanço da ciência e da razão. Sobre essa discussão Pombo (2003a, p.27) afirma que “Ontem como hoje, a enciclopédia continua construída, se não com base numa ideia de progresso perpétuo das luzes, pelo menos na de uma progressão exponencial dos conhecimentos”.

Entretanto, a modernidade reforçou em seu arcabouço o humanismo (teoria do pensamento comum ao movimento renascentista e iluminista), em especial do século XVIII, a visão antropocêntrica, isto é, o homem como centro do mundo, e também a visão que contempla e reforça a cisão ontológica do homem x máquina, nas ciências (mas em outros domínios, com outros predicados de oposição), traço desconstruído na pós-modernidade, essencialmente em visões que contemplem os agenciamentos maquímicos (DELEUZE; GUATTARI, 1995-1997) e híbridos (LATOURETTE, 1994) do homem com as tecnologias.

Antes do percurso digital, algumas enciclopédias abandonaram, em grande medida após os meados do

século XX, a organização disciplinar e a ordem alfabética, e apostaram em configurações combinatórias mais complexas, como a “Tábua de relações” da Enciclopédia Universal (1968-1975, França), a “Rotação dos círculos” da Enciclopédia Britânica (1973-1974, Inglaterra, 15ª edição), e as “Zonas de leitura” da Enciclopédia Einaudi (1979-1984, Itália): “[...] estamos perante dispositivos combinatórios que anunciam a curiosidade lúdica que serve de guia à navegação” (POMBO, 2003a, p.12). No começo da década de 90 do século XX, as enciclopédias começam a ser publicadas em CD-ROM, ou seja, ganharam versões eletrônicas, como a Britânica, em 1993, e a Encarta, a enciclopédia multimídia completa, em 1994, editada pela Microsoft (POMBO, 2003a).

Da enciclopédia eletrônica, *off-line*, à digital/virtual, *on-line*, Pombo (2003a) afirmou, que o hipertexto é o limite ideal da enciclopédia. Após quase uma década, o hipertexto aperfeiçoou-se por meio da descrição semântica e dados vinculados (ou estruturados), a Wikipédia tomou corpo desde a sua criação, em 2001, por Jimmy Wales.

É nessa direção e no bojo dessas reflexões histórico-filosóficas que se coloca a enciclopédia pós-moderna, no contexto da *Web Semântica* e da *Web Pragmática*; também denominada *Semiótica* por Monteiro e Moura (2014); Pietarinen (2003), e *Web Ubíqua*, por Mills (2008).

A *Web 3.0*, ou *Semântica*, foi definida por Berners-Lee *et al.* (2001) como uma extensão da *Web* tradicionalmente conhecida, mas com conteúdo significativo estruturado. À época, diante da falta de visualização dessa *Web* no ciberespaço e críticas sobre esse nome, Berners-Lee *et al.* (2006) afirmam que na verdade a *Web Semântica* seria uma *Web* de dados e informação. Segundo a Figura 1, a *Web Semântica* está em pleno curso. *Semântica*, no contexto tecnológico, significa a construção de uma infraestrutura adequada para os agentes inteligentes percorrerem a *Web* para extrair o conhecimento sobre algo ou alguma coisa (entidades) e executarem ações complexas.

A *Web Pragmática*, para Schoop *et al.* (2006), é uma extensão da *Semântica*, ou seja, complementa-a a partir da legitimidade, da colaboração e dos discursos de uma comunidade. Nessa direção, Pietarinen (2003) propõe a *Web Semiótica*, que seria a união entre a

Semântica e a Pragmática, em que é possível visualizar postulado da teoria semiótica, especialmente a teoria dos interpretantes em que máquinas e humanos fazem parte do processo de interpretação no ciberespaço (MONTEIRO; MOURA, 2016).

Já para Mills (2008), a nova *Web* denomina-se ubíqua ou Social Semântica e se consolidará em um futuro muito próximo, com a convergência entre as tecnologias e serviços que possibilitem o crescimento da conectividade social, *Web Social* (2.0), e o crescimento do conhecimento e do raciocínio da *Web Semântica*. Para melhor entender ou visualizar esses conceitos, a Figura 1 apresenta os vários estágios da *Web*.

Percebe-se que a flecha em ascendência aperfeiçoa e traz para a *Web Pragmática* tecnologias das demais *webs*, como o Wiki, programa de produção compartilhada (neste artigo a Wikipédia), um dos símbolos da *Web Social*, e manifestações da *Web Semântica*, como mecanismos de busca, *Knowledge Databases*, mecanismos semânticos e agentes inteligentes, entre outros.

A partir da possibilidade de atribuição de semântica aos dados, quantidade crescente de dados estruturados é publicada como *Linked Open Data* (LOD) sob a forma de *Resource Description Framework* (RDF). O DBpedia, desenvolvido em 2007, criado na Universidade de Leipzig, com o objetivo de extrair dados estruturados do conteúdo da Wikipédia e disponibilizar para a *Web* (DBpedia, 2015), o Freebase (banco de dados de curadoria comunitária sobre pessoas, fatos e coisas conhecidas, criado pelo *Metaweb* e comprado pelo *Google*, em 2010) e o *Wikidata* (projeto da *Wikimedia Foundation* para *web* de dados, criado em 2012) são exemplos bem sucedidos de bancos de modelagem de dados RDF, cujo conhecimento é extraído da Wikipédia. Essas iniciativas permitiram transformar a *web* de documentos em uma *web* de entidades (ZHU; IGLESIAS, 2015).

A definição de enciclopédia e a noção de redes semânticas adotadas por Eco (2013) aproximam-se da enciclopédia semiótica e do conceito de “semiosfera” de Lotman (1996) (mundo de signos que todos os humanos vivem) e, por que não dizer, da *Web Semântica* e da *Pragmática* manifestas nas plataformas digitais do ciberespaço (Figura 1). Assim, o ciberespaço “[...] é com-

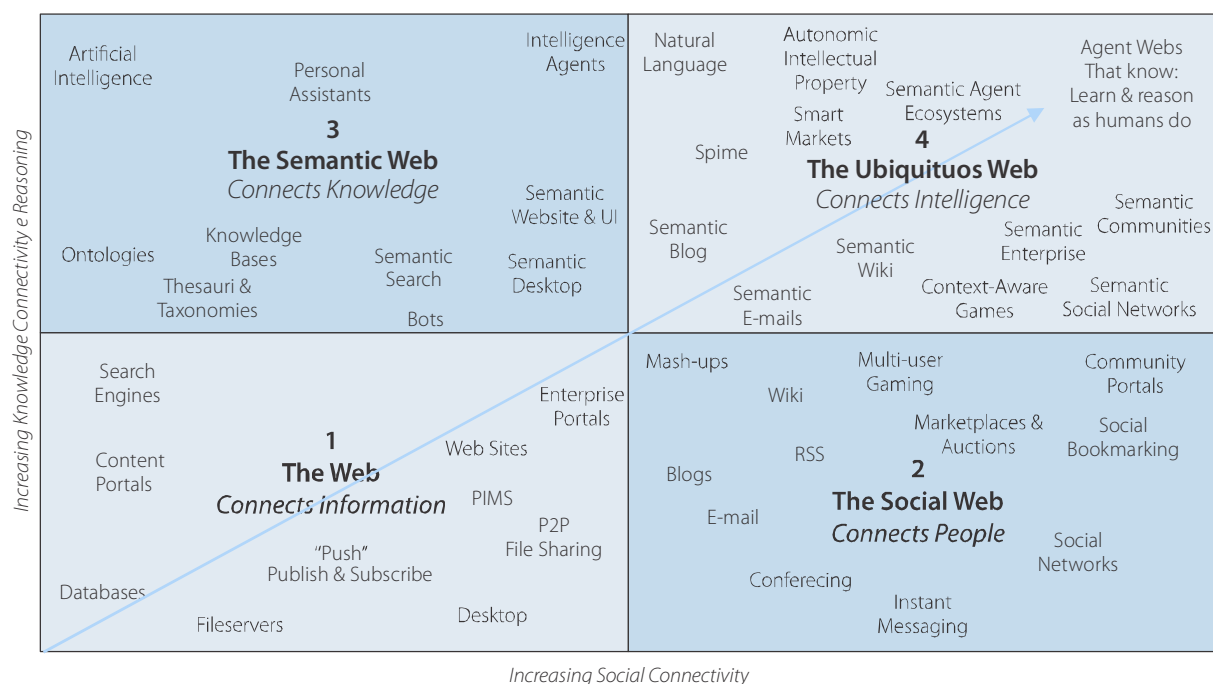


Figura 1. Qual a evolução da web para 2020?

Nota: PIMS: *Personal Information Managers*; RSS: *Really Simple Syndication*; P2P: *Point-to-Point*.

Fonte: Mills (2008, p.3).

preendido como território virtual de deslizamento de signos em tempo real, resultante do processo de digitalização da informação, esteio do capitalismo atual” (MAGOSSI, 2014, p.1).

Percebe-se a aparente tensão existente na produção e consumo de informação e conhecimento no ciberespaço. Se por um lado existe o fluxo livre de signos, a criação e defesa de *open source*, e o espírito do compartilhamento de conteúdo defendido pelos expoentes da computação na *Web Social* (por exemplo, Jimmy Wales da Wikipédia), por outro, essa situação gera lucros para empresas de indexação e busca da informação, como o *Google*. Transformando a afirmação de Pombo (2003a, p.12) em questionamento, será que “Mais uma vez a enciclopédia reflecte e simultaneamente traduz a situação dos saberes sua contemporânea. [?] Resta saber quais os desafios que lhe estão colocados”.

Índices e enciclopédias pós-modernas

Ao estudar a enciclopédia como organização semântica do saber e associá-la aos índices contemporâneos gerados pelos mecanismos de busca e, sobretudo, às complexas redes de conhecimento da *Web Semântica*, marcadamente coloca-se o principal atributo da pós-modernidade, isto é, a proeminência das tecnologias na sociedade atual e, por conseguinte, na produção, organização do conhecimento e processos de significação, sejam dos humanos, sejam das máquinas.

O ciberespaço como protomáquinas de virtuais possibilita a desterritorialização do universo simbólico e a formação de índices e enciclopédias virtuais na *Web*; os índices enciclopédicos podem estar mesclando esses dois conceitos. O fato é que as noções de índices e enciclopédias estão desestabilizadas e imbricadas pelo novo alfabeto digital e por esse espaço virtual de inscrição.

O hipertexto como ideia de enciclopédia, vislumbrado por Pombo (2003a, p.23) a partir da *Web*, a natureza enciclopédia da rede e o caráter remissivo da escrita digital fizeram surgir os serviços de indexação e busca, de forma que a autora afirma:

“[...] na continuidade dos desenvolvimentos recentes da enciclopédia, a rede faz mais do que

reunir e conservar a informação. O seu objectivo é também a indexação e a organização de todo o conhecimento mundial”.

Neste momento faz-se mister algumas reflexões sobre os índices. Para Mostafa (2013, p.90), “O índice assemelha-se mais às tabelas, pois indexar é relacionar [...]”. Acontece que, na Filosofia, toda relação implica um plano de consistência, pois pensar conceitos implica pensar os acontecimentos (e as relações) que lhes dão consistência.

Essas “empiricidades” ou acontecimentos fazem com que se denominem os mecanismos de busca como índices contemporâneos. Ao mesmo tempo em que se assume a indexação como meio de organização de conteúdos no ciberespaço, explicitando a relação com a área da Ciência da Informação, também se assume que esses índices são de outra ordem ou grau: evidenciam os agenciamentos maquínicos e híbridos característicos da contemporaneidade.

Para Bentes (2008, p.22):

Os ‘indexes’ são uma espécie de discursos construídos ou pelos humanos ou através dos *softwares* informáticos. Conforme o sentido da palavra, os ‘indexes’ propõem-se a indicar ou apontar a direção que pode ser tomada como o objetivo de encontrar e ter acesso a informações.

Índice, para a Ciência da Informação, é ao mesmo tempo um instrumento de organização e recuperação de informação. Se os índices na modernidade são definidos como “Relação de palavras ou frases, ordenadas segundo determinado critério, que localiza e remete para as informações contidas num texto” (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2004, p.1), na contemporaneidade eles podem ser definidos como “Signo que se encontra em relação de contiguidade temporal ou espacial com o objeto que refere” (INFOPÉDIA, 2013), ou seja, passam a ter uma acepção semiótica, mesmo que seja na ordem da informação. Os índices (dos mecanismos) podem ser definidos como uma enorme base de dados de informações importantes a respeito de *sites* na *Web*, mas, para completar a inteligibilidade do conceito, algumas definições interessantes de mecanismos de busca são apresentadas, por exemplo, os índices são povoados com etiquetas, outro tipo de metadados (BATTELLE, 2006); Soares utiliza a metáfora dos oráculos pós-modernos para os mecanismos (2008).

Já o *Google* é, para Levy (2012, p.15), “[...] uma máquina; um colecionador onívoro de informações”; “[...] um leiloeiro infalível”; “[...] um estudante de línguas, comportamentos e desejos assustadoramente hábil”; ainda, o *Google* é “[...] uma gigantesca máquina de aprendizagem baseada em inteligência artificial”.

O *Knowledge Graph* (KG), ou gráfico do conhecimento que aparece no lado direito dos resultados, é uma semantização da busca, realizada por agentes inteligentes, desenvolvido e implementado pelo *Google*, em dezembro de 2012, no domínio de língua portuguesa. Aparentemente mais um recurso, é um forte investimento em Inteligência Artificial, no campo de Recuperação da Informação. De acordo com Monteiro e Moura (2014), o KG agrega as seguintes possibilidades: *Autosuggest* (*turning strings into things*), *Semantic Tags* (*adding meaning to content*), *Entity Collections* (*exploring related topics*), *Geosearch Collections* (*exploring local topics*) e *Topical Weblinks* (*link out to the rest web*).

Observa-se o desenvolvimento das tecnologias (agentes inteligentes e algoritmos) a serviço da Recuperação da Informação, em especial pelos mecanismos de busca, tornando os índices contemporâneos e a enciclopédia (e as bases de conhecimento deles construídas a partir da estruturação dos dados, como a *DBpedia*, o *Freebase* e a Wikipédia) instrumentos poderosos de informação e conhecimento, que alavancam, inclusive, o comércio eletrônico. No ciberespaço, os índices contemporâneos e a Wikipédia são um *corpus* em contínua renovação e expansão e estão interligados por metadados de representação semântica, que povoam as entidades dos bancos de conhecimento e são legíveis pelos agentes inteligentes. A evolução de uma palavra-chave retirada de um *corpus* verbal para definição, desambiguação, descrição, contexto e interpretação mostra que o “*Google* é um cofre hiperenciclopédico do conhecimento humano” (LEVY, 2012 p.15), aproximando-se da enciclopédia.

Certamente esse índice tem se utilizado da plataforma da *Web Semântica* para atribuir os significados em contexto. Interessante observar que uma das definições de *Web Semântica* é, justamente, a de *Unambiguous Web* (SIEGEL, 2009). Sem dúvida, trata-se de uma complexa e imbricada rede semântica que vem sendo construída, no ciberespaço, cuja alcunha já se encontra como “*teia de significados*” (*Web of meaning*), por Landis (2014).

Ou seja, inicia-se a “semantização” em larga escala no ciberespaço, estruturando a informação enciclopédica disponível e ao mesmo tempo fornecendo informações para a *Web* e para os buscadores com tecnologias próprias para capturar dados semânticos. Nesse sentido, acrescentem-se outros *Knowledge Databases* que são construídos e operam “por trás” dos mecanismos de busca, como o *Freebase* (GOOGLE, 2016), o *DPpedia* e o *Wikidata*.

De acordo com o *site*, o “*Wikidata* é um banco de dados vinculados [...]” ou ainda um banco de conhecimento livre que pode ser lido e editado por seres humanos e máquinas (WIKIDATA, 2015, *online*). Faz parte da *Wikimedia Foundation*, uma organização sem fins lucrativos com conteúdo livre. A Fundação possui vários projetos: o *Wikimedia Commons* para imagem e outros arquivos de mídia, o *Wikitionary*, que fornece informações lexicais de palavras, bem como definições e sinônimos, a *Wikidata*, especializada em dados estruturados, e a Wikipédia, o mais conhecido e popular dos projetos da fundação mencionada.

A Wikipédia, como mencionado, foi criada em 2001, é hoje o projeto mais difundido e um dos serviços mais utilizados no ciberespaço, contando com mais de 14 milhões de artigos em centenas de línguas e dialetos (WIKIPÉDIA, 2015).

Burke (2012) afirma que seu projeto original se chamaria “*Nupédia*”, com proposta mais tradicional, e que, em determinado momento, mudou o rumo em direção à filosofia do compartilhamento e do conhecimento livre. Além do aspecto de seu tamanho, Burke (2012, p.342) identifica uma nova característica, a “*reflexidade*”, isto é, a capacidade de refletir e agir sobre si mesma, a partir das constantes revisões:

A Wikipédia também se caracteriza por algo que, apesar do perigo do antropocentrismo, podemos chamar de ‘autocrítica’, representada pelos avisos sobre os riscos à saúde intelectual como ‘A neutralidade deste artigo é discutível’ ou ‘Este precisa de citações adicionais para corroborar. Por favor, ajude a melhorar este artigo acrescentando referências fidedignas. Materiais sem fonte podem ser questionados e removidos’.

Em suma, destacam-se esses exemplos de conhecimento enciclopédico, ou banco de conhecimento, ou

ainda banco de dados estruturados, pois eles refletem definições pragmáticas de índices contemporâneos, pois têm desenvolvido algoritmos e agentes inteligentes para agregar semântica à Wikipédia e à indexação e busca de conteúdo no ciberespaço.

Considerações finais

O processo de significação (ou 'semantização') está ocorrendo no ciberespaço, especialmente, por meio da Wikipédia e dos outros projetos da *Wikimedia Foundation*, como a *Wikidata*, e por meio da *Web Semântica* e das bases de conhecimento, como o *Freebase* e o *DBpedia*, para citar os mais importantes. Esse movimento está construindo uma rede ou teia de significações para ser recuperada, ativada e interpretada por máquinas e humanos.

No artigo seminal sobre a *Web Semântica*, Berners-Lee *et al.* (2001) já afirmavam que seu verdadeiro poder será alcançado quando forem criados muitos programas que coletam o conteúdo da *web* a partir de diversas fontes de informação e agentes. A partir dessa premissa, pode-se atualmente visualizar a *Web Semântica* a partir da estruturação do conhecimento editado (socialmente) na Wikipédia, estruturado nos *Knowledge Databases*, capturado por agentes e visualizado pelo e nos índices contemporâneos como o *Google*, o primeiro a apresentar o KG. Isso significa que a *Web Semântica*, em larga escala, começa a ser percebida no ciberespaço por meio das tecnologias de visualização dos índices, como o supracitado, uma vez que os bancos de conhecimentos, apesar de disponíveis, necessitam de tecnologias para a coleta e visualização desse conteúdo.

Essa "semantização" terá impacto significativo no processo de atribuição do sentido, especialmente nos resultados de busca, pois "Isso nos afastará das pesquisas

por palavras-chave, pois a *Internet* deixará de ser um mundo de documentos para ser um mundo de dados que descrevem dados", frisa Santaella (2012, p.37) sobre a característica marcante da *Web 3.0*. Destarte, aproxima cada vez mais os índices das enciclopédias no processo de busca de conhecimento no ciberespaço.

Recordando Pombo (2002, *online*), "[...] a enciclopédia é um dispositivo semântico aberto, que remete para o mundo das coisas e dos acontecimentos [...]". Não obstante, essa Semântica, de alguma forma, se não se distancia da Linguística, aproxima-se indefectivelmente da Lógica (formal), da Engenharia, da Matemática e da Ciência da Computação. Essa semântica, os índices e a Wikipédia constroem, a partir de metadiscursos de referências e inferências, novas significações de tais construtos.

Talvez toda essa relação que se plasma nos resultados de busca de forma aparentemente orgânica, mas que tem a integração de várias fontes, especialmente da enciclopédia e de *Knowledge Databases*, possa ser chamada de *MetaWeb*, apropriando-se de Santaella (2012), na medida em que se verá crescer rede de inteligências, sejam estas inteligências humanas, artificiais, pessoais, coletivas ou híbridas. A função constitutiva do enciclopedista, hoje, requer conhecimentos, não do humanista, mas do sujeito com formação capaz de traduzir o mundo para a linguagem digital e de programação – do polímata do mundo clássico e medieval, ao especialista da Idade Moderna, que na contemporaneidade cedeu lugar ao informata.

A Wikipédia ("conhecimento rápido") atualiza a Semântica, com os *Knowledge Databases*, os índices, as descrições e as referências, povoando o ciberespaço de metadados e formando uma rede de significações. Para terminar essas reflexões, lembrando Lyotard (1988), a enciclopédia de hoje já são os bancos de dados de conhecimento.

Referências

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBR 6034 – Índice – apresentação*. 2. ed. Rio de Janeiro: ABNT, 2004.

BATTELLE, J. *A busca: como o Google e seus competidores reinventaram os negócios e estão transformando nossas vidas*. Rio de Janeiro: Elsevier, 2006.

BATISTA, G.; RODRIGUES, R. A construção de identidade na "geração fitness" do Instagram: a representação do eu e do corpo no ciberespaço. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 37., Foz do Iguaçu, 2-5 set. 2014. *Anais...* Foz do Iguaçu: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2014. p. 1-15.

- BENTES, V. P. A contribuição peirciana para a representação indexal de imagens visuais. *Encontros Bibli: Revista Eletrônica de Biblioteconomia e Ciência da Informação*, n. 25, p. 15-35, 2008. Disponível em: <<https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2008v13n25p15/878>>. Acesso em: 2 fev. 2014.
- BERNERS-LEE, T. *et al.* The semantic web: A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. *Scientific American*, p. 1-18, 2001. Available from: <autor alterar o link noonnonoon 9584.pdf>. Cited: Aug. 18, 2014.
- BERNERS-LEE, T. *et al.* *The semantic web revisited*. 2006. Available from: <http://eprints.soton.ac.uk/262614/1/Semantic_Web_Revisted.pdf>. Cited: Aug. 18, 2014.
- BURKE, P. *Uma história social do conhecimento I: de Gutenberg a Diderot*. Rio de Janeiro: Zahar, 2003.
- BURKE, P. *Uma história social do conhecimento II: da Enciclopédia à Wikipedia*. Rio de Janeiro: Zahar, 2012.
- CASTRO, C. M. *A prática da pesquisa*. São Paulo: McGraw-Hill, 1977.
- DBPEDIA. *DBpedia about*. Leipzig: DBpedia, 2015. Available from: <<http://dbpediawww.informatik.uni-leipzig.de/>>. Cited: Jan. 8, 2015.
- ECO, U. *Da árvore ao labirinto: estudos históricos sobre o signo e a interpretação*. Rio de Janeiro: Record, 2013.
- DELEUZE, G.; GUATTARI, F. *Mil platôs: capitalismo e esquizofrenia*. São Paulo: Ed. 34. 1995-1997. v. 1-5.
- GOOGLE. *Freebase 2016*. Disponível em: <<https://plus.google.com/109936836907132434202/posts/3aYFVNf92A1>>. Acesso em: 17 maio 2015.
- INFOPÉDIA. *Índice*. Porto: Porto Ed., 2003-2013. Disponível em: <<http://www.infopedia.pt/lingua-portuguesa/%C3%A9ndice>>. Acesso em: 8 jul. 2013.
- FOUCAULT, M. *As palavras e as coisas*. São Paulo: Martins Fontes, 2007.
- LATOUR, B. *Jamais fomos modernos*. São Paulo: Ed. 34, 1994.
- LANDIS, C. A web of meaning: Linked open data resources on the web. *College & Research Libraries News*, v. 75, n. 95, p. 492-505, 2014. Available from: <<http://crln.acrl.org/content/75/9/492>>. Cited: June 10, 2015.
- LEVY, S. *Google: a biografia. Como o Google pensa, trabalha e molda nossas vidas*. São Paulo: Universo dos Livros, 2012.
- LOTMAN, Y. *La semiósfera*. La semiótica de la cultura. Madrid: Cátedra, 1996.
- LYOTARD, J-F. *A condição pós-moderna*. 3. ed. Rio de Janeiro: J. Olympio, 1988.
- MAGOSSI, P. G. Comunicação e o outro: as ritualidades do ciberespaço e a aceleração da vida cotidiana. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 37., Foz do Iguaçu, 2-5 set., 2014. *Anais...* Foz do Iguaçu: Sociedade Brasileira de Estudos Interdisciplinares da Comunicação, 2014. p. 1-14.
- MILLS, D. *Project10X's. Semantic Wave 2008 Report: industry roadmap to Web 3.0 & multibillion dollar market opportunities*. 2008. Available from: <http://project10x.com/blog_downloads/SWsummary.pdf>. Cited: Dec. 2, 2014.
- MONTEIRO, S. D.; MOURA, M. A. Knowledge Graph and "semantization" in cyberspace: A study of contemporary indexes. *Knowledge Organization*, v. 41, n. 6, p. 429-439, 2014.
- MONTEIRO, S. D.; MOURA, M. A. Knowledge Graph e Peirce: uma abordagem semiótica dos índices contemporâneos. In: TOMAÉL, M. I.; ENGUELMANN, A. R. A. *Fontes de Informação digital*. Londrina: Eduel, 2016.
- MOSTAFA, S. P. *Filosofia da diferença e a Ciência da Informação*. Rio de Janeiro: E-Papers, 2013.
- PIETARINEN, A. V. *The semantic + Pragmatic web = the semiotic web*. 2003. Available from: <http://www.iadis.net/dl/final_uploads/200302C149.pdf>. Cited: Feb. 28, 2015.
- POMBO, O. *O hipertexto como limite de ideia de enciclopédia*. 2003a. Disponível em: <<http://www.educ.fc.ul.pt/hyper/enciclopedia/presentefuturo.pdf>>. Acesso em: 8 maio 2016.
- POMBO, O. *Para uma história da ideia de enciclopédia*. Alguns exemplos. 2003b. Disponível em: <<http://www.educ.fc.ul.pt/hyper/enciclopedia/cap2p1/antclass.htm>>. Acesso em: 8 jan. 2015.
- POMBO, O. *Projeto enciclopedista*. 2002. Disponível em: <<http://www.educ.fc.ul.pt/hyper/enciclopedia/cap1p1/palavra.htm>>. Acesso em: 8 jan. 2015.
- SANTAELLA, L. A tecnocultura atual e suas tendências futuras. *Signo y Pensamiento*, v. 30, p. 30-43, 2012.
- SANTOS, J.F. *O que é pós-moderno*. São Paulo: Brasiliense, 2006.
- SCHOOP, M. *et al.* The pragmatic web: A manifesto. *Communications of the ACM*, v. 49, n. 5, p. 75-76, 2006.
- SIEGEL, D. *Pull: The power of Semantic Web to transform your business*. London: Portfolio Hardcover, 2009.
- SOARES, C. A. *Anatomia da busca*. 2008. Disponível em: <<http://blog.pontolit.com.br/autor/>>. Acesso em: 8 abr. 2009.
- VRANDEËIÆ, D.; KRÖTZSCH, M. Wikidata: A free collaborative knowledgebase. *Communications of the ACM*, v. 57, n. 10, p. 78-85, 2014.
- WIKIDATA. *Welcome to Wikidata*. 2015. Available from: <https://www.wikidata.org/wiki/Wikidata:Main_Page>. Cited: May 18, 2015.
- WIKIPEDIA. *Sobre a Wikipedia*. 2015. Disponível em: <http://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Sobre_a_Wikip%C3%A9dia>. Acesso em: 10 jun. 2015.
- ZHU, G.; IGLESIAS, C. A. *Sematch: Semantic Entity Search from Knowledge Graph*. 2015. Available from: <<http://km.aifb.kit.edu/ws/sumpre2015/paper4.pdf>>. Cited: May 19, 2015.

