



XVII Encontro Nacional de Pesquisa em Ciência da Informação (XVII ENANCIB)

GT 2 – Organização e Representação do Conhecimento

WEB SEMÂNTICA E WEB PRAGMÁTICA: discussão crítica sobre versionamento na web e limites conceituais

SEMANTIC WEB AND PRAGMATIC WEB: critical discussion of versioning on the web and conceptual limits

Denysson Axel Ribeiro Mota¹, Nair Yumiko Kobashi²

Modalidade da apresentação: Comunicação Oral

Resumo: Estudo comparativo das propostas da Web Semântica e da Web Pragmática, com base nas teorias da organização e recuperação de informações. Para isso, são abordados os conceitos de sintaxe, semântica e pragmática, assim como as possíveis classificações da Web. A discussão mostra que há limites para introduzir significação em sistemas de informação e que termos como semântica e pragmática requerem abordagens críticas. De fato, a operacionalização de conceitos semânticos e pragmáticos ainda está longe de ser realidade nos sistemas de informação contemporâneos, no contexto da www.

Palavras-chave: Organização do Conhecimento. Recuperação da Informação. Representação da Informação. Web Semântica. Web Pragmática.

Abstract: *Comparative study of the proposals of the Semantic Web and Pragmatic Web, based on the information organization and retrieval studies. For the development of this paper, the origins of the concepts of semantics and pragmatics, as well as the classifications of the Web, were discussed. The analysis allowed understanding the limits to introduce semantics and pragmatics in information systems and that terms such as semantic and pragmatic require critical approach. Indeed, the operationalization*

¹Universidade Federal do Cariri - UFCA

²Universidade de São Paulo - USP

of semantic and pragmatic concepts is still far from reality in contemporary information systems, in the context of www.

Keywords: *Knowledge Organization. Information Retrieval. Information Representation. Semantic Web. Pragmatic Web.*

1 INTRODUÇÃO

A proposta de criação de uma Web Pragmática surgiu da percepção de que algo faltaria à Web Semântica para promover efetiva interação com o usuário, em termos de recuperação de informação. Diversos autores (SINGH, 2002; DE MOOR, KEELER e RICHMOND, 2002; DI MAIO, 2008) apontam o principal problema da Web Semântica: a ausência de elementos que complementem os significados dos conteúdos publicados na Web, questionando, ainda, se a adição de elementos semânticos a esses conteúdos é suficiente para atingir o objetivo de criar uma Web significativa, ou Web de significados, como muitas vezes é chamada. Vale ressaltar que o termo semântica, utilizado no contexto da Web, remete à Semântica formal, que é o modo de garantir a realização de funções interpretativas por processos computacionais.

Para de Moor, Keeler e Richmond (2002), a Web Pragmática é uma Web em que processos pragmáticos essenciais são definidos e automatizados. A Pragmática é definida como o estudo da relação do signo com seus intérpretes, intérpretes estes inseridos em um contexto, que pode ser tanto situacional, como individual ou social, segundo Morris (1985) e Peirce (1977). Estes últimos são amplamente reconhecidos como os precursores conceituais da Web Pragmática (DE MOOR, KEELER e RICHMOND, 2002; DI MAIO, 2008). Dessa forma, é possível ver a Web Pragmática como um conjunto de tecnologias para criar representações de contextos na Web, de forma a aprimorar os processos de recuperação da informação.

A necessidade de se considerar contextos na recuperação de informações não é uma novidade. Os estudos de relevância na recuperação já levavam em consideração os contextos. Ingwersen e Järvelin (2005), por exemplo, ressaltam que a recuperação de informações é um processo de busca que ocorre dentro de contextos determinados por diversos elementos: social, organizacional e cultural, assim como contextos de natureza sistêmica. Para estes autores, toda informação é construída e interpretada com base em dois elementos contextuais: o modelo de mundo do agente (seu contexto social e cognitivo) e a mensagem identificada em um contexto. Porém, deve-se pensar que a mensagem também foi construída dentro de um contexto, por um autor, com um modelo de mundo próprio. Quanto menos informação o agente tem sobre o

contexto de produção e consumo de mensagens, maior a liberdade para interpretá-las. Isso ocorre de forma semelhante quando se considera a interpretação não apenas de uma fonte informacional, mas também de diferentes informações cujas fontes contextuais não estão acessíveis ao agente.

Enquanto em ambientes livres, lúdicos ou poéticos essa liberdade seja considerada desejável, ela pode ser fonte de ruídos em ambientes informacionais de domínios científicos. Em outras palavras, a indeterminação dos significados pode resultar em menor acurácia no processo de recuperação. Porém a simples representação de contextos não soluciona os problemas; é imperativo determinar o que pode ser considerado como contexto suficiente para a interpretação (INGWERSEN; JÄRVELIN, 2005). Isso é necessário porque insuficiente informação de contexto de produção e consumo de mensagens pode promover ambiguidade e, assim, levar à necessidade de introduzir maior quantidade de informação contextual, e assim indefinidamente.

Dentro dessa problemática, e com base na visão de contexto suficiente para a interpretação de mensagens, de Ingwersen e Järvelin (2005), pergunta-se: o que pode ser considerado informação contextual suficiente para representar informação significativa? Para entender este problema e começar a delinear uma possível representação de contextos que envolvam a produção e consumo de informações, mesmo em domínios restritos à documentação de domínios científicos, é necessário esclarecer a noção de contexto, como proposto em estudos pragmáticos.

2 PRAGMÁTICA E WEB PRAGMÁTICA

Como dito anteriormente, a Web Pragmática tem como um de seus objetivos definir e representar os elementos pragmáticos considerados essenciais para a compreensão e interpretação de mensagens. Porém, por serem conceitos de uso recente, tais como Web Semântica e Web Pragmática, em conjunção com outras noções que podem ser abordadas de diferentes formas, inclusive referentes aos termos Semântica e à Pragmática, poderão ocorrer ambiguidades que demandam sua melhor compreensão.

Desde sua criação, a Web opera como uma plataforma de interação e comunicação entre indivíduos. Originalmente pensado para a comunicação entre bases militares, a Web vem sendo utilizada efetivamente entre universidades e acadêmicos, vem sendo continuamente ampliada para uso pela população em geral. Diante de seu contexto evolutivo, é possível afirmar que os elementos da comunicação escrita – Sintaxe, Semântica e Pragmática – já ocorrem na Web, desde seu surgimento, embora os elementos semânticos e pragmáticos tenham presença limitada

nesse ambiente. Com efeito, as pesquisas para tornar a Web mais significativa têm centrado esforços em projetos para adicionar elementos semânticos e pragmáticos aos conteúdos disponibilizados, com o objetivo de tornar os computadores aptos a realizar inferências complexas. Esses esforços deram origem a duas novas áreas de estudos e aplicação: a Pragmática Virtual, que estuda as ações da teoria da ação comunicativa e dos jogos de linguagem na Web (GRACIOSO, 2010), e a Web Pragmática que busca representar elementos pragmáticos para a Web.

Antes de aprofundar esta discussão terminológica, é necessário esclarecer os diferentes critérios que podem ser utilizados para caracterizar os tipos de Web. Um deles, definido por O'Reilly (2005), classifica a internet em estática ou interativa, esta última chamada de Web 2.0. Semelhante afirmação foi feita em face da percepção da Web como uma plataforma de acesso a serviços comerciais com características únicas, sendo citados como exemplos paradigmáticos as “eternas” versões beta, situações em que os serviços, por serem constantemente modificados, raramente são considerados versões definitivas ou completas.

Outra possibilidade de classificação da Web tem origem nas propostas de Berners-Lee, Hendler e Lassila (2001) que se caracteriza como uma pesquisa e um conjunto de procedimentos para adicionar significados aos conteúdos da Web. No entanto, o conceito de significado, desta proposta, não tem em vista, prioritariamente, o ser humano, mas a “compreensão” e a comunicação entre computadores. Essa característica promove a diferenciação entre Web Sintática e Web Semântica: na primeira, o computador é capaz apenas de identificar os diferentes elementos que compõem o conteúdo, definidos como uma sequência ou conjunto de caracteres, ou signos; na segunda, é possível inferir significado dos diferentes elementos do conteúdo. A Web Pragmática, por sua vez, amplia o conceito de significação proposto pela Web Semântica.

Nesta discussão, é importante ressaltar algumas questões: primeiro, que as classificações acima são realizadas sob distintas perspectivas: a classificação de O'Reilly tem foco na comunicação humana mediada por computador, ampliada tecnologicamente, mas considera fortemente a perspectiva mercadológica da Web; a de Berners-Lee, Hendler e Lassila é voltada para a interação do computador com os conteúdos disponíveis na Web, especificamente para as formas de representação de informações e de conhecimentos neste ambiente. Sob essa ótica, introduz novos conceitos e métodos para que o computador possa “entender” conteúdos e, assim, promover melhor recuperação de informações. Em segundo lugar, a proposta de O'Reilly foi apenas uma tentativa de dar nome a um fenômeno que já estava acontecendo, uma “simples” constatação da realidade, enquanto a de Berners-Lee, Hendler e Lassila cria uma nova Web,

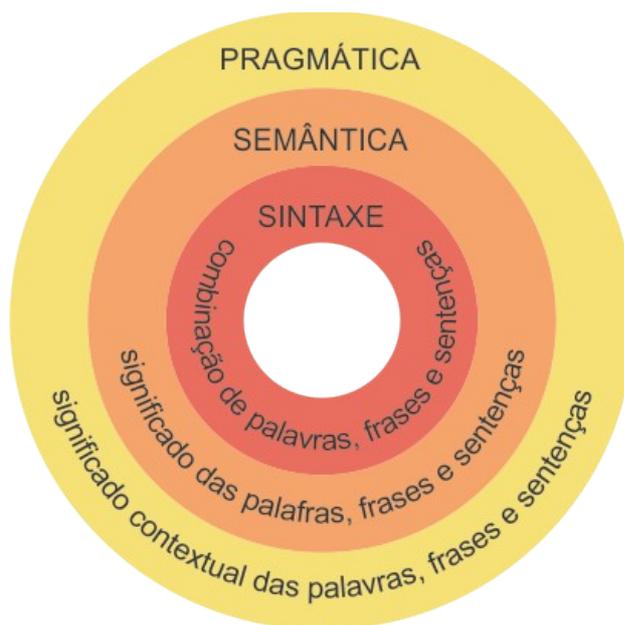
com novas possibilidades e ferramentas operacionais.

Por outro lado, ao considerar a Web, desde seu surgimento, como um ambiente criado para a comunicação, há a possibilidade de serem estudados elementos sintáticos, semânticos e pragmáticos neste ambiente, como dito anteriormente. A Pragmática, enquanto domínio de estudos dedicados à construção social de significados, ganhou na Web um campo de novas pesquisas e experimentações, sendo possível identificar e analisar diferentes formas de comunicação, dentro e entre os grupos sociais. No entanto, para diferenciar Web semântica de Web pragmática é necessário compreender melhor os conceitos e métodos previstos nesses modelos.

2.1 SINTAXE, SEMÂNTICA E PRAGMÁTICA

Sintaxe, Semântica e Pragmática são conceitos definidos segundo diferentes paradigmas. Nos estudos linguístico-semióticos de Morris (1985), ponto de partida conceitual deste trabalho, há três elementos que influenciam o significado das palavras: a sintaxe, a semântica e a pragmática (Figura 1). Os estudos de Morris, assim como os de vários outros pesquisadores, se baseiam nos trabalhos semióticos de Peirce (1977), principalmente na tríade representâmen-objeto-interpretante. Para Morris (1985) a Sintaxe é apenas a verificação da composição sígnica das palavras, a junção de letras que definem o significado, e que indicariam as diferenças dos significados de “rato” e “fato”, por exemplo.

Figura 1 – Sintaxe, semântica e pragmática



Fonte: Mota, 2015.

A Semântica, por outro lado, acrescenta à composição do significado da palavra sua posição em uma frase, no aspecto lógico-linguístico, e sua relação com seu *denotatum*, no aspecto semiótico-filosófico (TAMBA-MECZ, 2006; MORRIS, 1985).

Para Morris (1985, p. 55, tradução nossa), a Semântica é o estudo da “relação dos signos com seus designata e, com isso, com os objetos que podem denotar ou que, de fato, denotam”³. Esta visão é puramente semiótica e relaciona cada palavra ou termo com a imagem mental do objeto ideal que este designa, independentemente de outros elementos que possam influenciar a designação. Os estudos semânticos se constituíram de diferentes formas, de acordo com a época em que foram realizados ou com a base teórica adotada, porém, a forma mais clara e diferenciada das outras formas de estudo dos sentidos (sintático ou pragmático) é a que explora o significado derivado das regras de composição das sentenças, segundo estruturas, de certa forma pré-definidas, independentemente do contexto de uso (TAMBA-MECZ, 2006).

Dito de outro modo, a Semântica, no aspecto lógico-linguístico, concentra-se “apenas nas características linguísticas das significações, sem preocupação com seus demais aspectos, filosóficos ou psicológicos” (TAMBA-MECZ, 2006, p. 11). Busca-se, nessa perspectiva, identificar o significado literal das palavras e frases.

Segundo Bréal (1900), o significado não apenas deriva de regras impostas sobre as palavras, mas outros elementos, como a posição das palavras em uma frase, podem modificá-lo. Esta ordem, por exemplo, está definida nas regras de formação das frases, chamadas pelo autor de lógica da linguagem. Segundo Bréal, são essas regras que, entre outras coisas, não permitem que um adjetivo esteja numa flexão de gênero diferente da do substantivo.

Um alerta, porém, é feito por Cançado (2008) para o fato de que o estudo semântico das expressões, seja sob a visão linguística ou semiótica, que ignora o contexto de uso das expressões, é insuficiente para compreender de forma mais ampla como ocorre o processo de interpretação. Bréal (1900, p. 106) já apontava para o aspecto contextual da interpretação quando afirmou que a “expressão se adapta à coisa através das circunstâncias, o lugar, o momento, e a óbvia intenção do discurso”⁴.

O campo que estuda o uso da língua dentro de contextos é a Pragmática. Na Pragmática, elementos extralinguísticos são utilizados para complementar o significado das palavras, de modo a explicitar a intencionalidade do emissor e do receptor, dando maior importância à

³ No original: “relación de los signos con sus designata y, por ello, con los objetos que pueden denotar o que, de hecho, denotan”.

⁴ No original: “expression adapt itself to the thing through the circumstances, the place, the moment, and the obvious intention of the discourse”.

relação dos signos com seus usuários (MORRIS, 1985). Para Bréal (1900), o receptor tem metade da responsabilidade na interpretação (sendo a outra metade, responsabilidade do autor) e denomina de *subjective-element* o fator de interpretação derivado de elementos contextuais. Com os estudos do campo da Pragmática é possível identificar a intenção do emissor da mensagem, mesmo que as palavras, isoladamente, indiquem o contrário (PALMER, 1981).

Para Armengaud (2006), Morris é o primeiro autor a usar especificamente o termo Pragmática, embora os estudos deste campo já ocorressem desde a década de 40, com Bar-Hillel, e com especial destaque para o ano de 1950, quando foi publicado o artigo de Strawson “*On Referring*”, que abordava o uso de expressões dentro de contextos. Armengaud (2006) afirma que não há um único fundador da Pragmática, mas há vários: Peirce e Morris, são considerados fundadores diretos; Frege e Wittgenstein, são tidos como fundadores indiretos, graças à contribuição para a Semântica, do primeiro, e o estabelecimento do Paradigma da Comunicabilidade, do segundo; Carnap e Bar-Hillel, são fundadores intermediários, devido à sua influência nos trabalhos posteriores.

Todo sistema de interpretação é baseado no signo e, assim, a Linguística está dentro do escopo da semiótica (PEIRCE, 1977). Para este autor, o sentido é função do uso, e quando as frases são analisadas sem considerar sua ocorrência em contextos, haverá um sentido oriundo apenas das palavras e de seu encadeamento, porém, quando consideradas a identidade ou situações do falante e do ouvinte, o sentido é modificado, enriquecido (ARMENGAUD, 2006).

Morris (1985), por sua vez, estabelece, a partir dos estudos de Peirce, os níveis de semiose, de interpretação do signo e suas relações. O primeiro é o nível sintático, que aborda a relação dos signos entre si, o segundo, é o nível semântico, que relaciona os signos com os objetos e, o terceiro, o nível pragmático, que relaciona os signos com seus intérpretes. Para Morris, este último nível tem os seres vivos como intérpretes, e por isso, caracteriza também os aspectos bióticos da semiose, sejam eles psicológicos, biológicos ou sociológicos (ARMENGAUD, 2006).

Segundo Armengaud (2006), é Frege (2009) quem introduz a diferenciação entre sentido e referência. A diferença entre esses dois conceitos está diretamente relacionada aos planos em que atuam: o sentido está no plano interpretativo e, a referência, no plano extralinguístico. O exemplo clássico exposto por Armengaud (2006) é o caso das frases “a estrela da tarde” e “a estrela da manhã”, que têm sentidos diferentes, embora tenham a mesma referência: Vênus. Frege (2009) também determina que o fato de ser possível inferir sentido de uma frase não indica que sua referência tenha sido identificada. Cita como exemplo “o corpo celeste mais distante da terra” (p. 133) para mostrar que, apesar desta frase ter sentido, não há objeto de

referência para ela.

Para determinar o sentido, então, não basta conhecer a referência – se é que ela existe –, mas sim compreender o contexto de uso da palavra, que Frege denomina de princípio da contextualidade (ARMENGAUD, 2006). Outra contribuição de Frege, segundo Armengaud (2006), é a diferenciação entre linguagem científica e linguagem natural, em que a primeira tem a necessidade funcional de ser unívoca, marcando com precisão suas articulações, mas, a segunda, é polissêmica, ou seja, em sua riqueza e ambiguidade, promove maior poder afetivo e retórico.

Wittgenstein (1968, p. 41), de certa forma, relativiza as fundamentações semânticas ao dizer que “Você não pode prescrever a um símbolo o que lhe é permitido expressar. Tudo o que um símbolo *pode* expressar lhe é *permitido*” (grifos no original), afirmando com isso que a relação entre o signo e o seu significado não seria fixa ou pré-determinada. Posteriormente, ele será mais brando, afirmando que “toda palavra tem um significado. Este significado é atribuído à palavra. Ele é o objeto que a palavra designa” (WITTGENSTEIN, 2009, p. 15), e afirma que essa construção ou atribuição de significado às palavras é feita pelo uso da linguagem.

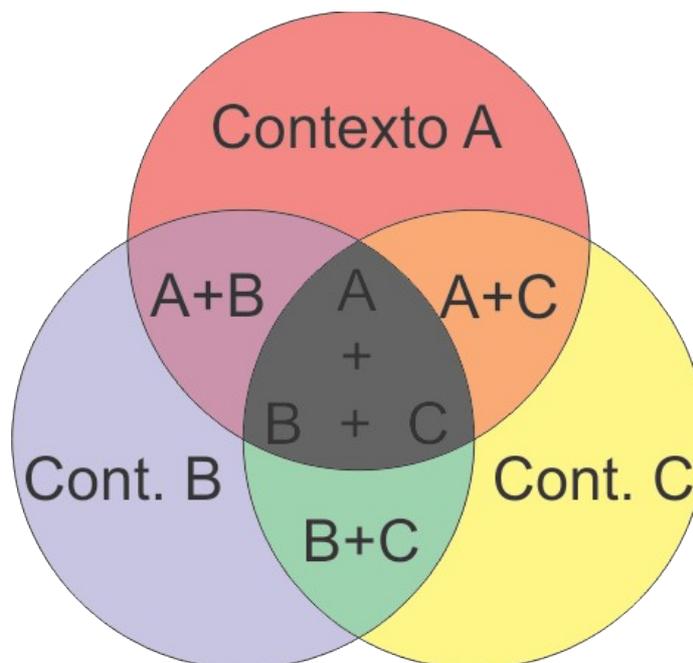
Para Blair, um estudioso da pragmática na Ciência da Informação, Wittgenstein é um dos principais responsáveis (ou o principal responsável) pela "virada linguística" ocorrida no século XX (BLAIR 2003, p. 9). Segundo este autor, os trabalhos de Wittgenstein evidenciaram que não se trabalha com ideias, mas com a descrição delas, e que muitos dos problemas filosóficos não eram realmente filosóficos, mas apenas decorrentes do uso inadequado da linguagem. Wittgenstein, conforme Blair (2003), mudou a forma de percebermos o pensamento ao mostrar que não pensamos em "ideias" ou "significados" e depois os traduzimos em palavras, mas que são as próprias palavras que vêm à mente, sendo a linguagem o próprio "veículo do pensamento" (WITTGENSTEIN, 2009, p. 146). Portanto, para Blair (2003), o ser humano não utiliza a linguagem apenas para expressar a ideia pensada, para traduzir o conhecimento, mas a utiliza para pensar.

O uso da linguagem e a construção de significados nasce, segundo Wittgenstein (2009), de forma social, entre pelo menos dois indivíduos: um instrutor e um instruído, um mestre e um discípulo. Dessa forma, a construção do significado ocorre de forma prática, isto é, é no uso da palavra que ela adquire significado. Toda interpretação é, nessa medida, um processo vinculado à ação, denominada de ação comunicativa. As diferentes ocasiões e propósitos guiam a construção de sentido dos enunciados, processo que o autor denomina de “jogo de linguagem” (WITTGENSTEIN, 2009), por se tratar de uma atividade social, partilhada e regulada. Esses “jogos de linguagem” seriam ilimitados, assim como os contextos de interpretação.

2.2 CONTEXTO

Nesta mesma linha, os contextos, além de serem ilimitados, não são únicos ou isolados, mas podem se “sobrepor”, criando, dessa forma, novos e diferentes contextos (Figura 2). O contexto é o cerne da Pragmática, um objeto de pesquisa intrigante para as diferentes áreas do conhecimento. O significado de contexto tem um histórico de divergências entre as diversas áreas que definem o conceito (VALLET *et al*, 2006), não havendo, portanto, um único significado aceito para toda situação ou tipo de informação veiculada (EDMONDS, 1999). Devido a essa diversidade, busca-se aqui reunir definições oriundas de diferentes áreas para identificar uma perspectiva adequada às propostas da Web Pragmática.

Figura 2 – Possibilidade de sobreposição de contextos



Fonte: Mota, 2015.

De forma geral, e no campo específico da Linguística, o contexto é composto por vários elementos (sociais, situacionais e/ou pessoais) que interferem no comportamento linguístico. As possibilidades contextuais são, de certa forma, infinitas, principalmente se forem considerados os sujeitos em situações que envolvem diferentes culturas, idiomas, domínios do saber e formação. Nessa mesma perspectiva, Armengaud (2006) define contexto como as situações comunicativas concretas que ocorrem em situações determinadas por lugar, tempo e identidade dos falantes. Tudo o que for necessário para avaliar o significado das elocuições é, então, contexto. O sentido das palavras deve ser apreendido a partir do sentido das frases onde elas

figuram e das frases dentro de um parágrafo, e assim consecutivamente. Para compreender, então, o que o enunciatório deseja informar, há que situar a palavra em seu contexto de uso na frase.

O contexto, para Lyons (1979), pode ser definido a partir de quatro elementos:

[...] situação espaciotemporal que inclui o falante e ouvinte, as ações que eles realizam no momento e vários objetos e eventos externo. [...] O ouvinte [...] interpreta esses [...] elementos corretamente mediante referência às características relevantes da situação. [...] o conhecimento partilhado pelo falante e pelo ouvinte do que se disse antes, na medida em que isso seja pertinente para a compreensão do enunciado. [...] a aceitação tácita pelo falante e ouvinte de todas as convenções, crenças e pressuposições relevantes 'dadas como certas' pelos membros da comunidade linguística a que o falante e o ouvinte pertencem (LYONS, 1975, p. 413)⁵.

Percebe-se que as possibilidades de uso de expressões linguísticas são bastante intrincadas sendo, em geral, pouco consideradas nos processos de recuperação de informações não por ser algo ignorado ou desconhecido, mas porque sua adequada representação implica levar em conta elementos que muitos desenvolvedores de sistemas de informações julgam demasiado complexos para serem operacionalizados.

É importante ressaltar que o conceito de contexto pode ser considerado ambíguo ou demasiado amplo e, por isso, alguns autores utilizam outros termos para designar o conjunto de circunstâncias nas quais o sujeito se encontra durante o processo de busca de informações. Alguns exemplos dessas delimitações são as de Julien e Michels (2002), que utilizam o termo "situação", e Allen and Kim (2001), que preferem utilizar o termo "tarefa". Por outro lado, Ingwersen e Järvelin tratam o contexto como um conjunto de fatores, uma série de elementos determinados por itens como "a tarefa de uma pessoa, sua fase e situação" (INGWERSEN; JÄRVELIN, 2005, p. 1, tradução livre). Além disso, outros indivíduos e componentes também podem influenciar o processo.

Armengaud (2006, p. 82) afirma que o “conceito de contexto adquire vigor e consistência ao se aproximar da noção de mundo possível”, restringindo-se, assim, as possibilidades contextuais a um conjunto de situações, denominado de *conjunto-contexto*, que

⁵ No original: [...] spatiotemporal situation which includes the speaker and hearer, the actions they are performing at the time and various external objects and events. [...] The hearer [...] interprets these [...] elements correctly by reference to the relevant features of the situation. [...] the knowledge shared by the speaker and hearer of what has been said earlier, in so far as this is pertinent to the understanding of the utterance. [...] the tacit acceptance by the speaker and hearer of all the relevant conventions, beliefs and presuppositions 'taken for granted' by the members of the speech-community to which the speaker and hearer belong.

são as mais prováveis de ocorrer. O conceito de mundos possíveis é compartilhado por Eco (2000), que exemplifica com a palavra baleia. Para Eco, existem dois mundos possíveis: um onde ‘*baleia*’ é interpretado como um peixe, e outro onde é interpretado como mamífero. Ambas as interpretações podem estar corretas porque podem pertencer a diferentes concepções de mundo.

Considerando essas variações, e a abordagem da questão por diferentes autores, é perceptível que alguns elementos podem ser representados de forma relativamente fácil e outros não, sendo este um dos principais problemas presentes em estudos de usuários. Porém, para avaliar inicialmente as várias possibilidades de pesquisa, será considerado aqui o contexto de forma mais ampla, com seus vários possíveis elementos, de acordo com as definições de Armengaud (2006) e de Ingwersen e Järvelin (2005).

Na perspectiva pragmática, Teun Van Dijk (1977a) trabalha com a noção de *frame*. Este é definido como um conjunto de conhecimentos de mundo (sejam eles conhecimento técnico, de idioma, posicional, situacional etc.) que possibilita elaborar diferentes interpretações das sentenças. Assim, diferentes *frames* permitem interpretar uma mesma frase de diferentes formas.

Eco (1987) aborda os *frames* como uma “estrutura de dados que serve para representar uma situação estereotipada”⁶ (p. 114, tradução nossa). Essas estruturas conteriam um conjunto de informações que ditam as ações e reações do sujeito na situação em que este se encontra. Dessa forma, se as informações do *frame* se confirmam, o sujeito sabe como deve agir, porém, caso existam informações divergentes no *frame*, o sujeito entra em um estado ou condição em que é necessário adaptar-se à situação para agir de acordo com as equivalências situacionais armazenadas como conhecimento prévio. Nestas situações, o sujeito modifica o *frame* em que ocorreram as divergências ou cria um novo para a situação específica.

A representação de contextos, em forma de *frames*, pode ser considerada como puramente representação do conhecimento, pois, ao contrário do que ocorre com as representações de registros informacionais, o que se busca explicitar aqui são os diferentes conhecimentos situacionais que influenciam a interpretação textual. O autor afirma que já houve iniciativas para explicitar o conhecimento de mundo, porém, o sucesso dessas experiências foi limitado devido à complexidade e variação dos contextos possíveis de ocorrer.

Edmonds (1999), em consonância com Eco (1987) e Van Dijk (1977a, 1977b), afirma que, nos sistemas naturais, um único efeito pode ter várias causas e que, para sua adequada representação, é mister ignorar alguns elementos para que a ampla quantidade de fatores não

⁶ No original: “estructura de datos que sirve para representar una situación estereotipada”.

gere demasiada complexidade ao realizar a modelagem do mundo possível, ou do domínio desejado. Considerando o efeito como a interpretação de uma palavra ou sentença, as causas podem ser várias, e elas variarão de acordo com elementos psicológicos e sociais construídos desde a infância. Assim, no âmbito dos sistemas de informação, para a adequada representação de informações contextualizadas, é necessário delimitar os contextos, considerando apenas os mais plausíveis ou prováveis. Dito de outro modo, os conceitos de *mundo-possível*, *frame* e *conjunto-contexto* tornam operacionalizável a inclusão de contextos em sistemas de informação.

Para Borlund (2003), o contexto é a percepção do usuário em situação de busca de informação, portanto, nada mais é que um construto psicológico que representa, internamente, o conhecimento e as suposições sobre o mundo em um determinado momento e sua necessidade informacional dentro desse mundo. De forma semelhante, um autor também tem seu contexto, pois, ao representar seu conhecimento, ele expõe ali uma série de elementos de contexto que o circundam, começando pelo tema escolhido e a forma de representação.

A classificação de contextos em interno e externo (físico, social e biológico) é proposta por Edmonds (1999). No primeiro, o contexto interno é o próprio conhecimento do indivíduo (ou grupo social), utilizado no processo de aprendizagem mediante a (re)lembração e o reconhecimento de diferentes situações para melhor se adequar ao presente; já o contexto externo é caracterizado pelo ambiente compartilhado por um grupo, chamado também pelo autor de “*context-as-a-resource*”⁷ e “*context-we-inhabit*”⁸, que pode ser, por exemplo, um contexto temporal. Mas sua delimitação, sem considerar o contexto interno, é complexa, principalmente porque o contexto interno pode se referir a um grupo de pessoas, e separar o que seria interno e externo para esse grupo traz claras dificuldades.

Na segunda classificação, o contexto externo é subdividido em três outros contextos: físico, social e biológico. O contexto físico é constituído pelos elementos espaciais e temporais que um grupo de pessoas pode compartilhar, e, para o autor, são os mais simples ou claros de serem analisados; o contexto social é o que gera, segundo Edmonds (1999), a maior quantidade de elementos em comum entre os indivíduos, principalmente por serem oriundos da cultura e da linguagem, mas, ao mesmo tempo, é demasiado rico, ou complexo, para ser analisado; já o contexto biológico são as experiências emocionais, de consciência ou de conhecimento compartilhado que um grupo de indivíduos pode ter. Eles são extremamente importantes no processo de aprendizagem, mas seu estudo é complexo pois, muitas vezes, são confundidos ou ocultados pelos contratos sociais.

⁷ Tradução livre: contexto como um recurso.

⁸ Tradução livre: contexto que habitamos.

Por outro lado, Armengaud (2006) nota que Carnap estabelece distinção entre dois tipos de dependências entre palavra e significado: uma dependência não-essencial, quando o significado de uma palavra ou frase é constituído pelas palavras/frases que a precedem (um contexto interno ao texto); a outra, chamada de essencial, quando o significado varia de acordo com elementos extralinguísticos (um contexto externo ao texto). São semelhantes aos “ditos e não ditos” de Eco (1984, p. 97).

Nesses estudos, a dependência não-essencial é completamente voltada para o texto, sendo realizada pelo leitor no momento de uso da informação. Estes aspectos têm sido utilizados em sistemas de recuperação de informações e de indexação/catalogação semiautomáticos, como na identificação de palavras relevantes dentro de frases e de frases dentro de parágrafos. Essa identificação ainda não é perfeita (LANCASTER, 1993), porém, é realizada com grande velocidade e tem resultados considerados aceitáveis.

Mesmo que Armengaud (2006) afirme que a primeira dependência se preocupa apenas com o texto, é possível perceber que ela também se relaciona com o contexto extralinguístico do autor. As escolhas das palavras para expressar as ideias, autores para citar ou, até mesmo, o uso de voz passiva ou ativa é uma influência de elementos extralinguísticos.

A segunda, a dependência essencial, no entanto, é mais complexa, pois é dependente dos indivíduos que interagem com o texto e o contexto da interação, sejam eles histórico, acadêmico ou social. Essa dependência é especialmente relevante quando considerados os processos de recuperação de informações, em que a leitura ocorre de forma diagonal, visto que se procura, basicamente, identificar palavras/frases que indiquem relevância para o usuário, ou por consulta ao resumo do texto. Neste caso, o contexto é, especificamente, a necessidade de informações, demandadas pela tarefa do indivíduo, sendo também influenciada por seu histórico profissional ou acadêmico.

Como dito anteriormente, Julien e Michels (2002), Allen and Kim (2001) e Ingwersen e Järvelin (2005), abordam essa dependência de perspectivas variadas, cada um com diferentes objetivos e resultados, o que mostra a complexidade dos estudos de contextos nas ações humanas.

3 WEB SEMÂNTICA E WEB PRAGMÁTICA: limites conceituais

Ao discutir os conceitos de semântica, pragmática e contexto, assim como suas aplicações no ambiente Web há, certamente, limites em relação ao que pode ser considerado efetivamente como conteúdo semântico ou aspecto pragmático. Estas discussões não são novas, pois ocorrem desde o surgimento do conceito de Web Semântica. Porém, é necessário esclarecer

alguns pontos que podem parecer conflitantes ou incompletos.

O que é questionável é até que ponto a Web Semântica, e sua sucessora, a Web Pragmática, interpretam os conceitos de semântica e pragmática.

O que se percebe, ao serem analisados os exemplos e explicações presentes na literatura consultada, e nos dados abertos disponíveis na web, é que os aspectos considerados semânticos se resumem, via de regra, à inclusão de maior quantidade de metadados padronizados para descrever os acervos informacionais; pode-se afirmar que não há, de fato, adição de significados, no sentido proposto pelos estudos semânticos, ou de adição de contextos de interpretação das informações, como sugere a Pragmática. Ao se considerar que os dados e, neste caso, os metadados, são expressos mediante sintaxe, então é possível afirmar que não se conseguiu, ainda, representar informações com elementos semânticos ou pragmáticos no ambiente Web, como propõem e afirmam as propostas da Web Semântica e da Web Pragmática.

Não há, então, a possibilidade de encontrar elementos semântica ou pragmaticamente representados, em toda a amplitude e profundidade conceitual proposta por áreas como a Semiologia e a Semiótica, na Web Semântica e na Web Pragmática. Como afirma Gärdenfors (2004), a Web Semântica não é ontológica, pois a descrição dos elementos de conteúdo informacional, na rede, é parcial e não cobre as possibilidades essenciais que uma ontologia realmente poderia prover. O que se tem nas novas propostas para a Web, tanto no caso da Web semântica, quanto na Web pragmática, são aplicações simplificadas, expressas por linguagem lógica formal, ao ambiente digital voltado para a comunicação entre dispositivos computacionais.

Almeida, Souza e Fonseca (2011) apresentam um conjunto de definições sobre a Semântica, no contexto da Web Semântica. Afirmam que, dentro deste escopo, a abordagem semântica que mais se aplica é a semântica formal. Os autores de textos da área, no entanto, não deixam isso claro, ou não a definem de forma adequada. Gärdenfors (2004), por sua vez, critica a visão da computação de que toda semântica poderia ser reduzida à lógica de primeira ordem, ou teoria dos conjuntos. Para este autor, o raciocínio humano é complexo, não sendo possível formalizar essa complexidade por meio dos métodos da Semântica Lógica.

O que o computador realiza, nestes ambientes, são manipulações de símbolos e de proposições dadas como verdadeiras. Não é possível afirmar que o computador é dotado de inteligência ou que seja capaz de realizar inferências complexas, de forma semelhante ao Argumento da Sala Chinesa, de Searle (1980), no qual se questiona se um processamento de dados ou símbolos realizado de forma predeterminada realmente indica inteligência ou autonomia pela entidade processadora.

É possível perceber, nos textos apresentados e discutidos acima, que dentro dos estudos e vertentes atuais da Semântica há elementos passíveis de aplicação à Web Semântica, principalmente ao se considerar a forma como os seres humanos processam os signos e os compreendem. Da mesma forma, é possível pensar na ampliação e aprofundamento de aplicação das propostas semióticas à Web Pragmática. Mas, no momento atual, a interação e a negociação de significados, como postos nos trabalhos de Wittgenstein (1968, 1994) e Morris (1985), não são realidade no ambiente da Web Semântica e da Web Pragmática.

Dito isto, é importante destacar que os aspectos das teorias pragmáticas para aplicação no ambiente Web, são de operacionalização complexa e as teorias, métodos, procedimentos e tecnologias atualmente disponíveis não estão suficientemente desenvolvidos para suportar esses tipos de aplicações.

Reitera-se, portanto, que a possibilidade de representação de informação com o uso de elementos específicos das teorias pragmáticas para promover o aprimoramento das formas de representação e recuperação de informações em ambientes informacionais digitais, diante das limitações teóricas, metodológicas e tecnológicas e conceituais presentes, estão distantes de uma possível implementação, em curto prazo.

4 CONSIDERAÇÕES FINAIS

Constata-se, pela discussão acima apresentadas sobre as questões semânticas e pragmáticas, que as propostas da Web Semântica e da Web Pragmática estão longe de cumprir as promessas que vêm sendo feitas. As limitações tecnológicas atuais permitem pensar que o que de fato é possível adicionar aos sistemas de recuperação de informações (SRI), hoje, é uma quantidade maior de dados de natureza sintática aos metadados já existentes nas páginas web. Com isso, mesmo que se seja possível simular operações semânticas em computadores, com o uso de proposições lógicas simples, ou operações pragmáticas, com a adição de elementos contextuais, com a finalidade de realizar seleções mais precisas de informações, não é possível afirmar que tais operações signifiquem a criação de uma Web Semântica ou mesmo de uma Web Pragmática.

Quando consideradas essas limitações, a representação de informações tanto em RDF quanto em XTM, muitas vezes consideradas como linguagens com propósitos semânticos, são, de fato, linguagens para adição de dados sintáticos. Deve-se observar, porém, que estas linguagens e padrões promovem avanços nas formas de representar informação para recuperação. Eles permitem que os dispositivos tecnológicos troquem dados de forma facilitada, colaborando, de certa forma, tanto para qualificar melhor os dados armazenados na rede, quanto

para promover a interoperabilidade entre diferentes tipos de sistemas.

REFERÊNCIAS

ALLEN, B.; KIM, K.-S.; Person and context in information seeking: Interactions between cognitive and task variables. **The New Review of Information Behaviour Research**, v. 2, p. 1–16, 2001.

ALMEIDA, M.; SOUZA, R.; FONSECA, F. Semantics in the Semantic Web: A Critical Evaluation. **Knowledge Organization**, Berlin, v. 38, No.3, pp. 187 – 203, maio/jun. 2011.

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**. New York: Scientific American, 2001. Disponível em: <<http://www.sciam.com/2001/0501issue/0501berners-lee.html>>. Acesso em: 09 out. 2016.

BLAIR, D. C. Information Retrieval and the Philosophy of Language. **Annual Review of Information Science and Technology**, v. 37, pp. 3-50, Medford, 2003.

BORLUND, P. The Concept of Relevance in IR. **Journal of the American Society for Information Science and Technology**, v.54, p. 913-925, 2003.

BOUNDLESS. **Boundless Psychology**. 201X. Disponível em <<https://www.boundless.com/psychology/textbooks/boundless-psychology-textbook/>> Acesso em: 09 ago. 2016.

BRÉAL, M. **Semantics**: studies in the science of meaning. New York: Henry Holt & Company, 1900.

CANÇADO, M. **Manual de Semântica**: noções básicas e exercícios. Belo Horizonte: UFMG, 2008.

DE MOOR, A.; KEELER, M.; RICHMOND, G. Towards a pragmatic web, In: UTA, Priss et al. Conceptual Structures: Integration and Interfaces. **Lecture Notes in Computer Science**, v. 2393, p. 235-249, 2002. Disponível em <<http://www.cspeirce.com/menu/library/aboutcsp/richmond/web.pdf>>. Acesso em 23 mar. 2016.

DI MAIO, P. The Missing Pragmatic Link in the Semantic Web. **Business Intelligence Advisory Service Executive Update**. v. 8, n. 7, 2008.

ECO, U. **Lector in Fabula**: la cooperación interpretativa en el texto narrativo. Barcelona: Lumen, 1987.

ECO, U. **O conceito de texto**. São Paulo: T. A. Q. /EDUSP, 1984.

ECO, U. **Os limites da interpretação**. São Paulo: Pioneira, 2000.

EDMONDS, B. The Pragmatic Roots of Context. In: PROC. OF THE 2ND INTERNATIONAL AND INTERDISCIPLINARY CONFERENCE ON MODELING AND USING CONTEXT. Berlin; Heidelberg; New York, v. 1688, 1999. **Anais...** v. 1688, p. 119-132, 1999.

FREGE, G. **Lógica e Filosofia da Linguagem**. São Paulo: EDUSP, 2009.

GÄRDENFORS, P. Conceptual Spaces as a Framework for Knowledge Representation. **Mind and Matter**. v. 2, n. 2, pp. 9–27, 2004. Disponível em <http://www.cc.gatech.edu/classes/AY2013/cs7601_spring/papers/gaerdenfors.pdf>. Acesso em 14 jul. 2016.

GRACIOSO, L. de S. Justificação e a ação de informação no contexto da pragmática virtual. **Liinc em Revista**, Rio de Janeiro, v.6, n.2, p. 286- 300, set. 2010.

INGWERSEN, P.; JÄRVELIN, K. **The turn**: integration of information seeking and retrieval in context. Dordrecht, The Netherlands: Springer, 2005.

JULIEN, H.; MICHELS, D. Intra-individual information behaviour in daily life. **Information Processing and Management**, v. 40, p. 547-562, 2004.

LANCASTER, F. W. **Indexação e resumos**: teoria e prática. 2. ed. Brasília: Briquet de Lemos, 2004.

LYONS, J. **Introdução à Linguística teórica**. São Paulo: Nacional, Edusp, 1979.

LYONS, J. **Introduction to Theoretical Linguistics**. Cambridge: Cambridge University Press, 1975.

MORRIS, C. **Fundamentos de la teoría de los signos**. Barcelona: Paidós, 1985.

MOTA, D. **Representação e recuperação de informação em acervos digitais nos contextos da web semântica e web pragmática**: um estudo crítico. 2015. Tese (Doutorado em Cultura e Informação) - Escola de Comunicações e Artes, Universidade de São Paulo, São Paulo, 2015. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/27/27151/tde-27012016-135403/>>. Acesso em 14 jul. 2016.

O’ REILLY, T. **What is Web 2.0**. Disponível em: <<http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>>. Acesso em 14 jul. 2016.

PALMER, F. R. **Semantics**: a new outline. Cambridge: Cambridge University Press, 1981.

PEIRCE, C. S. **Semiótica**. São Paulo: Perspectiva, 1977.

SEARLE, J. Minds, Brains and Programs. **Behavioral and Brain Sciences**. v. 3, pp. 417–57, 1980. Disponível em <<http://purl.org/spar/cito/>>. Acesso em: 14 jul. 2016.

SINGH, M. P. **The Pragmatic Web**: Preliminary thoughts. 2002. Disponível em: <<http://lstdis.cs.uga.edu/SemNSF/Singh-Position.pdf>>. Acesso em: 14 jul. 2016.

TAMBA-MECZ, I. **A Semântica**. São Paulo: Parábola Editorial, 2006.

VALLET, D. et al. **Personalized Information Retrieval in Context**. 2006. Disponível em <

<http://ir.ii.uam.es/~acemedia/publications/mrc06.pdf>>. Acesso em 20 jul. 2016.

VAN DIJK, T. A. **Cognição, discurso e interação**. São Paulo: Contexto, 1992.

VAN DIJK, T. A. Context and cognition: Knowledge frames and speech act comprehension. **Journal of Pragmatics**, v. 1, p. 211–232, 1977a.

VAN DIJK, T. A. **Text and Context**: Exploration in the semantics and pragmatics of discourse. London: Longman, 1977b.

WITTGENSTEIN, L. **Investigações Filosóficas**. Petrópolis: Editora Vozes, 2009.

WITTGENSTEIN, L. **Tractatus Logico-Philosophicus**. São Paulo: Companhia Editora Nacional/EDUSP, 1968.