

Ciência da Informação, Ciência da Computação e Recuperação da Informação: algumas considerações sobre os métodos e tecnologias da informação utilizados ao longo do tempo

**Information Science, Computer Science, And Information Retrieval: some considerations
about the methods and technologies of information used along time**

Rachel Cristina Vesu Alves
Doutoranda do Programa de Pós-graduação em Ciência da Informação da Faculdade de
Filosofia e Ciências da UNESP – Campus de Marília
rachelvesu@yahoo.com.br

Vângela Tatiana M. Banhos
Mestranda do Programa de Pós-graduação em Ciência da Informação da Faculdade de
Filosofia e Ciências da UNESP – Marília
vangela@marilia.unesp.br

Ana Lúcia A.O. Bichei
Mestranda do Programa de Pós-graduação em Ciência da Informação da Faculdade de
Filosofia e Ciências da UNESP – Marília
donanalucia@yahoo.com.br

José Eduardo Santarem Segundo
Tecnólogo em Processamento de Dados.
Doutorando do Programa de Pós-graduação em Ciência da Informação da Faculdade de
Filosofia e Ciências da UNESP – Campus de Marília
santarem@marilia.unesp.br

WOIDA, Luana Maia
Mestranda do Programa de Pós-graduação em Ciência da Informação da Faculdade de
Filosofia e Ciências da UNESP – Campus de Marília
woida@marilia.unesp.br

RESUMO

A recuperação da informação vem sendo estudada sob duas visões: da área da Ciência da Informação e da Ciência da Computação. A discussão fundamenta-se em uma breve revisão de literatura, buscando contemplar os precursores das temáticas de tratamento informacional (organização e representação), recuperação da informação, bem como, as Tecnologias de Informação e Comunicação. Com o objetivo de estabelecer as relações e contribuições entre as áreas de Ciência da Informação e Ciência da Computação, em relação à recuperação da informação, destaca-se uma contribuição mútua entre as áreas, no intuito de garantir acessibilidade e usabilidade ótimas dos ambientes informacionais diversos.

Palavras-Chave: Ciência da Informação. Ciência da Computação. Recuperação da Informação. Tecnologia de Informação e Comunicação. Organização Informacional.

ABSTRACT

The information retrieval has been studied under two visions: from the Information Science area, and the Computer Science area. The discussion has its fundamentals on a short literature revision, trying to contemplate the precursors of the informational treatment thematic (organization and representation), information retrieval, as well as the Information and Communication Technologies. Having the goal of establishing the relations and contributions between the Information Science and Computer Science areas, relating to the information retrieval, a mutual contribution between the areas stands out, with the objective of guaranteeing the optimal accessibility and usability of the diverse informational environments.

Keywords: Information Science. Computer Science. Information Retrieval. Information and Communication Technology. Informational Organization.

1 INTRODUÇÃO

A preocupação com a recuperação da informação vem de épocas remotas e passou a ser amplamente abordada e discutida em diversas áreas do conhecimento. Ganhou atualmente maior destaque devido à quantidade crescente de informações disponibilizadas e pelo uso de Tecnologias de Informação e Comunicação – TIC's, constituindo-se em um tema comum entre áreas tais como a Ciência da Computação e Ciência da Informação - CI.

A CI é uma área que busca investigar, entre outros assuntos, a geração, a organização, o processamento, a recuperação e o uso da informação, em diversos ambientes. É, por natureza, uma área interdisciplinar, por isso seus estudos recebem contribuições de várias áreas, ao mesmo tempo em que seus produtos são subsídios para as outras áreas do conhecimento. Além disso, está intimamente ligada com as Tecnologias de Informação e Comunicação vigentes.

Deste modo, é nítido que a CI estabelece uma relação com a Ciência da Computação no que se refere à recuperação da informação. Assim, a discussão está fundamentada em uma breve revisão de literatura da CI, buscando contemplar os precursores das temáticas de organização e representação da informação, da recuperação da informação, bem como das Tecnologias de Informação e Comunicação.

Lembrando que a preocupação com a recuperação da informação é uma constante na CI e também sendo contemplada nos estudos de Ciência da Computação, o objetivo a ser tratado está em estabelecer as relações e contribuições da CI para a área de Ciência da Computação e vice-versa, no que se refere à recuperação da informação.

Tendo em vista que a organização, representação e recuperação da informação sempre estiveram atreladas ao uso das tecnologias vigentes, propõe-se abordar os princípios e metodologias de organização e representação da informação estabelecidos na CI, bem como as ferramentas e

TIC's da área de Ciência da Computação para facilitar os processos de recuperação, as quais procuramos enfocar na Web Semântica.

Para isso, ter-se-á como base de análise as idealizações de Paul Otlet com a questão da organização do conhecimento; as idealizações de Vannevar Bush na questão das tecnologias para a recuperação da informação; e o atual estudo da Web Semântica proposto por Tim Berners-Lee, que busca uma melhor recuperação baseada não só em tecnologias avançadas, mas também com uma melhor organização e representação do conteúdo disponibilizado.

2 PRECURSORES E FATOS HISTÓRICOS DA CIÊNCIA DA INFORMAÇÃO: ALGUNS ASPECTOS SOBRE RELAÇÃO COM A TECNOLOGIA DA INFORMAÇÃO

O desenvolvimento humano tem proporcionado inovações tecnológicas que contribuíram para que todas as áreas do conhecimento, tanto teórica como prática, se desenvolvessem mais rapidamente e com resultados mais expressivos. Neste sentido, as novas tecnologias criadas nas últimas décadas, em especial as TIC's, tornaram-se ainda mais evidentes nos últimos anos.

Seus reflexos podem ser percebidos na CI por ser uma área intimamente ligada às tecnologias de informação. Vários teóricos, tal como Saracevic [1996], consideram a tecnologia como elemento chave no desenvolvimento desta ciência, pois a criação de ferramentas tecnológicas subsidia o desenvolvimento das teorias no propósito de alcançar os objetivos estabelecidos na CI em relação aos problemas aos quais ela se dedica resolver. Contudo é importante ressaltar algumas características da CI, tal como sua origem, características e relações com as tecnologias de informação, para que possam ser criadas bases para uma posterior discussão.

Pode-se dizer que a CI é uma ciência recente, sua origem é marcada pela revolução científica e tecnológica após o período da Segunda Guerra Mundial [Messias, 2005]. A origem da CI, de acordo com Pinheiro [1999, p. 175], advém de suas raízes,

[...] de um lado a Documentação e, de outro a recuperação da informação. Na primeira o que importa é o registro do conhecimento científico, a memória intelectual da civilização e, no segundo, as tecnologias de informação. Ciência e Tecnologia foram os elementos fertilizadores e propulsores de seu nascimento, fruto do crescimento de equipes científicas, do aumento do número de cientistas e pesquisadores, e da aceleração de pesquisas, portanto, de conhecimento, além dos desenvolvimentos tecnológicos, esforços decorrentes sobretudo da 2ª Guerra Mundial. E as tecnologias, principalmente os computadores a fazem emergir.

Uma definição freqüentemente retomada na literatura é a do teórico Borko. Este, por sua vez, fundamentou-se nas discussões e primeiras tentativas de diferenciar a CI tanto da Biblioteconomia como da Documentação, ocorridas no *Georgiatech* entre 1961 e 1962.

Ciência da Informação é a disciplina que investiga as propriedades e o comportamento da informação, as forças que governam seu fluxo de informação, e os meios de processar a

informação para otimizar sua acessibilidade e uso. Esta é relacionada ao corpo de conhecimento relativo à origem, coleta, organização, armazenamento, recuperação, interpretação, transmissão, transformação e utilização da informação [...] [Borko, 1968, tradução nossa].

Há que se considerar esta uma das definições mais importantes, pois corrobora com as afirmações de que a área é interdisciplinar e que tem influências de outras áreas sob o prisma pelo qual percebe os problemas relacionados à informação, bem como todas as formas de trabalho realizado sob o ciclo informacional. Nesse contexto, a abordagem sobre as tecnologias da informação torna-se de particular interesse para os objetivos da presente discussão, uma vez que não se trata apenas de novidades de informática, mas sim das principais ferramentas tecnológicas e métodos que venham a contribuir para o tratamento da informação.

Pode-se dizer que a preocupação em organizar e tornar acessível por meio da recuperação os registros do conhecimento se constitui um tema antigo na CI. Isso pode ser notado nas primeiras tentativas do homem de armazenar, organizar e recuperar a informação registrada, sendo que em cada época utilizava-se das tecnologias disponíveis para que a recuperação da informação pudesse ser realizada. Essa vertente da origem da CI é utilizada por alguns autores, tal como Shera apud Messias [2005].

Ainda em relação à origem, destacam-se alguns autores que estão diretamente ligados ao desenvolvimento e uso tecnológico na CI em seus respectivos tempos, autores conhecidos e que tem papel decisivo na história da área; são eles o belga Paul Otlet e o americano Vannevar Bush. Atualmente destaca-se o inglês Tim Berners-Lee, que será abordado nessa discussão no intuito de confrontar os estudos anteriores com os atuais. Estes teóricos utilizaram, direta ou indiretamente, as tecnologias de informação e métodos de organização e representação da informação no desenvolvimento de suas pesquisas, contribuindo fundamentalmente para o desenvolvimento das áreas de CI e Ciência da Computação.

O cientista belga Paul Otlet, reconhecido como o “pai” da documentação, representa um marco dentro da história relativa dos precursores da CI. A visão ampla de Otlet revolucionou o modo de trabalhar com a informação em sua época, pois não só sistematizou as questões teóricas presentes no problemático crescimento exponencial dos documentos, mas principalmente estabeleceu metodologias e técnicas na tentativa de solucionar a questão da recuperação, acesso e circulação da informação [Santos, 2006]. Otlet fez-se valer das tecnologias disponíveis na época para que pudesse colocar em prática suas idealizações de universalização do acesso ao conhecimento¹. Juntamente com Henri La Fontaine, Otlet foi líder do Movimento Bibliográfico

ocorrido na Europa no final do século XIX e início do XX, no qual procurou desenvolver meios de acesso ao conhecimento. Seus estudos foram de grande contribuição para,

[...] a ampliação do conceito de documento; o estabelecimento de sistemas de tratamento e recuperação da informação como são entendidos atualmente; a estruturação de redes internacionais de cooperação para a coleta e disseminação da informação e o estabelecimento da Documentação [Santos, 2006, p. 14].

Contudo, é no estabelecimento de sistemas de tratamento e recuperação da informação que se concentra seu trabalho. Otlet e outros pesquisadores envolvidos fundaram em 1895 o Instituto Internacional de Bibliografia – IIB e iniciaram uma revolução nesta área. Ao contrário dos bibliotecários, basicamente centrados no documento, Otlet buscou desenvolver técnicas e métodos de organização baseado no conteúdo dos documentos.

Pretendiam uma organização mundial da informação e para isso suas pesquisas voltaram-se para o desenvolvimento de um sistema de informação especializado e de produtos, tais como o Repertório Bibliográfico Universal que contemplava a produção científica da época. Outra importante ferramenta desenvolvida por Otlet em 1899 foi a Classificação Decimal Universal - CDU, técnica que ainda é utilizada por grande parte das bibliotecas do mundo. Para Otlet a CDU era uma representação sistemática do conhecimento que, por meio de códigos padronizados, estabelecia os relacionamentos e vínculos entre os assuntos representados e posteriormente com os objetos. Deste modo, pode-se dizer que a grande contribuição de Otlet foi o enfoque dado ao conteúdo informacional e às técnicas e métodos de tratamento informacional desenvolvidas para a disponibilização, acesso e recuperação da informação.

Em fins da Segunda Guerra Mundial destacou-se Vannevar Bush [1945], cientista americano que coordenou mais de 6000 cientistas, bem como a aplicação da ciência para o desenvolvimento de sistemas armamentistas. Contribuiu para avaliar o estado da ciência e da tecnologia e preocupou-se em tornar o conhecimento humano mais acessível. Fez comentários que se identificariam diretamente com os interesses da futura CI e sua interdisciplinaridade com profissionais dedicados à tecnologia da informação.

Diagnosticou a dificuldade cada vez maior do desenvolvimento científico em função de ineficientes sistemas de produção, organização, acesso, recuperação e disseminação da informação devido ao crescimento do conhecimento humano. Alertou para a perda de conteúdos importantíssimos em meio a uma imensidade de outros textos e conseqüente duplicação de pesquisas. Mas o problema não se restringia ao quantitativo aumento de informações, e sim na maneira de processá-la, na demanda de pesquisas e experiências que pediam processamentos que fizessem circular eficazmente um volume infinito de informação. Para gerenciar este problema, havia a necessidade de novas tecnologias (novas metodologias, instrumentos e máquinas).

Já naquela época Bush abordou preocupações atuais da comunidade científica e da humanidade como um todo: Qual a maneira eficiente de armazenar e recuperar o conhecimento humano? Como fazer uso apropriado do crescente volume de informações? Como impedir que uma informação relevante passe despercebida em meio a uma abundante geração de conhecimento? Como mecanizar o registro de idéias e experiências de forma a não estacionar no tempo em função da limitada capacidade de memória?

Bush explicava que a mente humana funciona por associação de idéias e elaboração de trilhas que levam uma idéia a outra, e afirmava que o homem não poderia duplicar artificialmente tal processo, mas poderia aprender com ele [Bush, 1945]. Ao sugerir a forma do pensar para transformar informação em conhecimento inseriu a noção de associação de palavras/conceitos na organização da informação. Acabou por descrever a idéia do que conhecemos hoje por hipertexto.

Mencionou que o registro feito por meio da escrita, fotografia e gravação com suportes e recursos que envolviam papel, filme, disco, lápis e máquina de escrever, se aperfeiçoaria e evoluiria, fazendo surgirem novos meios de transmissão do conhecimento. Expôs que o registro de idéias precisava, além de estar armazenado em algum lugar, ser consultado. Sugeriu a redução de espaço para disposição de informações contidas em uma enciclopédia e em todo um acervo de biblioteca, minimizando espaço, custo e aumentando a possibilidade de acesso [Bush, 1945].

Sua visão é imensamente importante para a CI, pois além dos aspectos já citados, falou em facilitar e agilizar o registro, a busca e recuperação da informação, alertando para o problema de localização da informação, ligado à superficialidade dos sistemas de indexação adotados (normalmente ordem alfabética ou numérica).

Bush idealizou um mecanismo para automatizar as ações de armazenar, indexar e recuperar conhecimento, chamando esse aparelho de Memex (Memory Extension). Foi uma das principais contribuições que levaram ao desenvolvimento teórico de um equipamento que tinha por princípio reproduzir a capacidade de associação das idéias humanas. Essa era a característica essencial do Memex, qual seja, a possibilidade de se estabelecer associação de idéias, ligações entre um item e outro. O Memex, caso chegasse a existir, seria uma máquina hipertexto, um precursor do computador pessoal, seria um misto de arquivo e biblioteca. Um dispositivo onde se armazenariam livros, publicações, registros, anotações e fotos, e que poderia ser consultado com extrema velocidade e flexibilidade, como se fosse uma extensão da memória humana.

Infelizmente a tecnologia da época não acompanhou o conhecimento teórico de Bush e assim não permitiu o desenvolvimento do Memex, porém é fácil perceber que sua idéia deu origem a relação hipertextual, utilizada hoje em dia em grande escala pela internet. O pensamento em forma de conexões é realmente algo diferenciado na época de Bush, visto que todas as teorias e

pensamentos eram construídos linearmente. Com tamanha visão e imaginação, pode-se afirmar que Bush foi um pioneiro de uma idéia que tantos anos depois originou a World Wide Web.

Apresenta-se, no quadro abaixo, um breve resumo cronológico com os estudos de teóricos e suas principais contribuições para a CI.

TEÓRICO(S)	ESTUDO(S)	CONTRIBUIÇÕES
Paul Otlet 1895-1934	Organizar a informação para posterior recuperação do conteúdo dos documentos.	- criou a Documentação; - Classificação Decimal Universal (CDU); e - repertório Bibliográfico Universal (RBU)
Vannevar Bush 1939-1945	Buscar soluções para o problema da explosão informacional.	- criou o analisador diferencial, primeiro calculador analógico completo. - arquitetou o Memex, um dispositivo para armazenamento e recuperação de todo o tipo de conteúdo informacional.
Tim Berners-Lee Década 90	Criador dos protocolos HTTP, URL, da linguagem Html que formam a tecnologia WWW.	Desenvolveu conceitos da Web Semântica

QUADRO 1: Quadro de teóricos da área de Ciência da Informação e Ciência da Computação, seus estudos e contribuições.

FONTE: Elaborado pelos autores.

Observa-se neste quadro que os estudos desde Otlet a Berners-Lee sustentam o objetivo de desenvolver um método/técnica de recuperação da informação que atenda ao usuário de maneira fácil, rápida e eficiente². As relações entre esses estudos serão abordadas a seguir.

3 VELHOS PRINCÍPIOS, NOVAS ABORDAGENS: o Repertório Bibliográfico Universal e a Web Semântica

Conforme já apontado a recuperação da informação é uma questão de interesse de duas áreas, a CI e a Ciência da Computação, é o ponto de convergência de técnicas, métodos e tecnologias para um fim comum.

Os avanços tecnológicos modificaram o modo de realizar as tarefas em diversos ramos de atividade, contudo, conforme aponta Alvarenga [2001] "Mudam-se os meios, sofisticam-se os instrumentos e surgem nomes novos para designar coisas velhas. Entretanto, a essência das coisas permanece". Em outras palavras, é preciso destacar que as atuais Tecnologias de Informação e Comunicação possuem princípios de organização, representação e recuperação da informação, os quais já foram pensados e até mesmo utilizados em outros contextos.

Podemos visualizar essa inter-relação quando tratamos, por exemplo, do Repertório Bibliográfico Universal e da Web Semântica. Duas tentativas de se oferecer acesso ao conhecimento mundial, tendo como características épocas e contextos distintos, porém, com alguns princípios semelhantes.

3.1 Repertório Bibliográfico Universal

O desenvolvimento de técnicas de organização e difusão da informação são ideais buscados por muitos estudiosos, tal como Paul Otlet, que se destaca por tentar reunir teoria e prática na criação de metodologias para análise e síntese do conhecimento, no intuito de viabilizar seu amplo acesso. Partindo desse ideal, Otlet se dedicou a projetos de socialização do conhecimento; como por exemplo, a Enciclopédia Documentária, o Mundaneum e o Repertório Bibliográfico Universal (RBU) ou ainda Bibliografia Universal; e teve como base o tratamento, análise e síntese do conhecimento registrado. Porém, o mais importante para o presente trabalho é o RBU, o primeiro repertório criado pela equipe de Otlet que se caracteriza como um inventário de todo o conhecimento disponibilizado.

Podemos considerar o RBU o primeiro sistema de recuperação da informação criado, sendo que “A metodologia empregada nos repertórios combinou três elementos: a) a Classificação Decimal Universal; b) as fichas 7,5x 12,5 e folhas soltas de tamanho padronizado; e c) o princípio monográfico” [Santos, 2006, p. 53]. Esses três elementos compõem os conceitos, técnicas e metodologias para a construção do sistema de recuperação da informação por meio dos Repertórios Bibliográficos. O sucesso do sistema de recuperação da informação elaborado por Otlet, por meio do RBU, não se deve somente a padronização adotada tanto no tamanho das fichas, mas também em seu preenchimento e principalmente nos métodos e técnicas adotados para o tratamento do conteúdo informacional [Santos, 2006]. Os métodos e técnicas adotados foram: o Princípio da monografia, o Princípio da continuidade e da pluralidade da elaboração, o Princípio da multiplicação dos dados.

O principal fator de êxito do RBU proposto por Otlet foi o Princípio da monografia, mais conhecido como Princípio monográfico, no qual o conteúdo informacional do documento era seccionado e separado em fichas ou folhas soltas padronizadas. Esse processo garantia a seleção das informações mais significativas para serem tratadas posteriormente pelos outros princípios. O Princípio de continuidade e pluralidade consistia na elaboração das fichas e folhas soltas em campos de dados padronizados, garantindo a inclusão de dados e informações úteis para a representação e

localização do documento. O Princípio da multiplicação dos dados consistia na multiplicação das fichas para que pudessem ser incluídas em diversos pontos de acesso do sistema [Santos, 2006].

Observa-se claramente que os princípios estabelecidos, juntamente com o uso da CDU, compõem uma estrutura para a representação das informações em um sistema de recuperação da informação. E apesar de incipientes quando comparadas com sistemas de recuperação da informação contemporâneos, os princípios elaborados por Otlet comportam um valor significativo, pois formam a base para os sistemas modernos.

Ainda que trouxessem contribuições para todas as áreas do conhecimento, os projetos de Otlet esbarravam em questões técnicas como a falta de ferramentas tecnológicas que otimizassem a recuperação da crescente massa informacional que estava surgindo. Assim, a efetivação das idealizações de Otlet só iria ocorrer com os posteriores avanços tecnológicos na área de informática e o desenvolvimento das TIC's. Embora muito dos princípios de tratamento informacional não tenham sido utilizados, vê-se hoje em dia uma retomada de velhos princípios, tais como os estabelecidos por Otlet e Bush, em uma nova abordagem marcada pelo avanço tecnológico atual.

3.2 Web Semântica

As TIC's originaram novos ambientes de disponibilização da informação, tal como a Web. Entretanto, ao mesmo tempo em que trouxeram soluções, intensificaram ainda mais o número de informações disponibilizadas e, conseqüentemente, os problemas de recuperação da informação. Contudo, presencia-se atualmente uma mudança de paradigma onde o conteúdo informacional passa a ser o foco e não mais o documento em si ou o suporte informacional no qual está veiculado [Robredo], 2004). Neste sentido, vêm sendo desenvolvidas novas ferramentas, técnicas e métodos para tentar melhorar a questão da recuperação da informação, um exemplo dessas propostas refere-se ao surgimento da Web Semântica.

Desenvolvida por Tim Berners-Lee e liderada pela W3C a Web Semântica é um projeto a ser implantado e vem sendo apontada como um caminho para solucionar os problemas de recuperação da informação na web, pois conta em sua estrutura de implantação com métodos mais adequados de tratamento informacional (representação e organização), bem como o estabelecimento do contexto semântico em que a informação se insere para que sejam desenvolvidas técnicas mais eficazes de recuperação [Alves, 2005].

O problema enfrentado com as atuais ferramentas de busca na web é que não são capazes de “raciocinar” sobre os dados e informações, como os seres humanos que fazem associações de

idéias. Partindo desse princípio a Web Semântica traz em seu escopo, ferramentas tecnológicas, como os agentes inteligentes, capazes de raciocinar e inferir sobre os dados, na tentativa de melhorar a recuperação. Mas para isso é preciso que o conteúdo informacional seja devidamente tratado, em outras palavras, a Web Semântica propõe uma nova forma de organização informacional para tornar o conteúdo dos recursos mais significativos aos raciocínios das máquinas [Alves, 2005; Santarem Segundo, 2004].

A estrutura da Web Semântica se assemelha aos sistemas de recuperação da informação, pois também trata do armazenamento, representação, organização e acesso aos recursos informacionais, [Souza; Alvarenga, 2004] com o diferencial de possuir nova estrutura tecnológica e organizacional que proporciona uma recuperação mais eficiente. A Web Semântica se constitui como um caminho na busca de soluções para os problemas já mencionados e é difícil prever seu futuro, pois promover a “compreensão” do significado das informações requer a implantação de estruturas complexas, tais como os agentes inteligentes, ainda em desenvolvimento na área de Inteligência Artificial [Alves, 2005].

A Web Semântica estabelece uma relação com a CI pela utilização das teorias de organização informacional, ou seja, todo embasamento teórico, filosófico, metodológico e conceitual da Web Semântica se relaciona ao núcleo duro da área [Souza; Alvarenga, 2004]. Além disso, a Web Semântica tem como base o uso de ferramentas tecnológicas para a implantação de sua estrutura e sob esse aspecto se relaciona com a Ciência da Computação, tal como os sistemas de recuperação da informação presentes na CI.

Como podemos perceber, a Web Semântica se constitui de duas formas, de um lado a estrutura tecnológica que viabiliza a recuperação da informação de modo mais ágil, e de outro a estrutura organizacional da informação que efetiva o funcionamento da recuperação, com base no tratamento do conteúdo informacional.

4 CONSIDERAÇÕES FINAIS

Analisando as propostas dos estudiosos abordados neste trabalho: Paul Otlet com a questão da organização do conhecimento; Vannevar Bush na questão das tecnologias para a recuperação da informação; e Tim Berners-Lee com o atual estudo da Web Semântica, pode-se verificar que essas propostas apresentam uma estreita relação no que se refere à recuperação da informação e conseqüentemente com a área de CI e Ciência da Computação. Mais do que isso, as propostas mostram que os princípios, métodos e tecnologias adotadas são comuns entre elas, porém, foram

estabelecidas em contextos diferentes devido ao avanço das tecnologias da informação. Em outras palavras, trata-se de uma nova abordagem para princípios já estabelecidos anteriormente.

A Web Semântica engloba muito dos princípios e idéias elaboradas por Vannevar Bush e Paul Otlet. A associação de idéias tal como a mente humana, elaborada por Bush, está contemplada na Web Semântica, que estabelece em sua proposta que a recuperação seja mais significativa por meio da associação de idéias entre conceitos a serem compreendidos pelos agentes inteligentes. Além disso, toda contribuição dada por Bush para o desenvolvimento das tecnologias, como a teoria de hipertexto, por exemplo, atua direta ou indiretamente na construção da Web Semântica, que pode ser considerada como um progresso a caminho das propostas idealizadas por Bush na questão de recuperação da informação.

Já em relação às idealizações de Otlet observa-se que a Web Semântica também se preocupa com o conteúdo do documento, tal como existiu no RBU. O foco no conteúdo permite na Web Semântica toda a aplicação da estrutura tecnológica necessária para seu estabelecimento e conseqüentemente, um tratamento informacional mais detalhado e adequado de acordo com os princípios que estabelece. Além disso, quando falamos em tratamento informacional, nota-se na Web Semântica alguns princípios estabelecido por Otlet, tal como o princípio monográfico, com o foco no conteúdo informacional e não no suporte ou formatação (representado na estrutura da Web Semântica pela linguagem XML - Extensible Markup Language); o princípio de continuidade e pluralidade, com a representação padronizada de informações e a inclusão de informações complementares ou adicionais (representado na estrutura da Web Semântica pelo uso de metadados); e o princípio de multiplicação, proporcionado pela representação e descrição ampla do recurso e a diversidade de pontos de acesso à eles (representado na estrutura da Web Semântica pelas ontologias e associações entre conceitos). A grande diferença é que os princípios já elaborados por Otlet se apresentam aperfeiçoados em uma nova abordagem tecnológica.

Em se tratando de contribuições da CI para a Ciência da Computação, bem como da Ciência da Computação para a CI, identifica-se, a partir do exposto trabalho que isso se dá mediante a recuperação da informação, com a aplicação das tecnologias da informação e da organização informacional. A Ciência da Computação fornece o meio, as ferramentas tecnológicas para o desenvolvimento do ambiente de recuperação da informação; já a CI fornece os métodos e técnicas de tratamento informacional (organização e representação) dos documentos disponíveis e faz uso das aplicações tecnológicas em seu fazer. A atuação conjunta das duas áreas favorece a constituição de uma estrutura tecnológica e informacional, que contempla as metodologias de tratamento informacional sedimentadas na área de CI e as aplicações tecnológicas desenvolvidas na área de Ciência da Computação. Sendo assim, uma área complementa a outra para que suas ferramentas e

métodos atuem em conjunto no intuito de garantir a acessibilidade e usabilidade ótima nos mais variados tipos de ambientes informacionais.

NOTAS:

(1) Otlet não distingue informação e conhecimento, mas utiliza o termo conhecimento para designar toda a informação registrada fruto de uma produção humana (um documento, que poderia ser um livro, ou um mapa, um boletim, um periódico etc). Neste trabalho nos referimos ao termo conhecimento, tal como utilizado por Otlet.

(2) Destacam-se ainda os seguintes autores que contribuíram para a recuperação da informação: Claude Elwood Shannon (1949), Calvin Mooers (1950-1951), Hans Peter Luhn (década de 50), Maron e Kuhns (década de 60), Gerald Salton (década de 70 e 80) e McGill (década de 80).

REFERÊNCIAS

ALVARENGA, L. A teoria do conceito revisada em conexão com ontologias e metadados no contexto das Bibliotecas tradicionais e digitais. *Data Gramma Zero – Revista de Ciência da Informação*, Rio de Janeiro, v. 2, n. 6, dez. 2001. Disponível em: http://www.dgzero.org/dez01/Art_05.htm. Acesso em: 31 jan. 2003.

ALVES, R. C. V. *Web Semântica: uma análise focada no uso de metadados*. 2005, 180f. Dissertação (Mestrado em Ciência da Informação) - Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2005.

BERNERS-LEE, T.; HENDER, J.; LASSILA, O. *The Semantic web: a new form of web content that is meaningful to computers will unleash a revolution of new possibilities*. [S. l.: S. n.], [2001?]. Disponível em: <http://www.scientificamerican.com>. Acesso em: 09 jan. 2005.

BORKO, H. Information Science: what is it? *American Documentation*, Washington, v. 19, n. 1, p. 3-5, Jan. 1968.

BUCKLAND, M. K. Information as thing. *Journal of the American Society for Information Science*, NEW York, v. 45, n. 5, p. 351-360, 1991.

BUSH, V. As we may think. *The Atlantic Monthly*, Boston, v. 176, n. 1, p. 101-108, 1945. Disponível em: <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>. Acesso em: 28 fev. 2004.

MESSIAS, L. C. da S. *Informação: um estudo exploratório do conceito em periódicos científicos brasileiros da área de Ciência da Informação*. 2005, 184f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências – UNESP, Universidade Estadual Paulista, Marília, 2005.

PINHEIRO, L. V. R. Campo interdisciplinar da ciência da informação: fronteiras remotas e recentes. IN: PINHEIRO, L. V. R. (Org.). *Ciência da informação, ciências sociais e interdisciplinaridade*. Brasília: IBICT; Rio de Janeiro: IBICT/DDI/DEP, 1999. p. 155-182.

ROBREDO, J. Organização dos documentos ou organização da informação: uma questão de escolha. *Data Gramma Zero – Revista de Ciência da Informação*, Rio de Janeiro, v. 5, n. 1, fev. 2004. Disponível em: http://www.dgzero.org/fev.04/Art_05.htm. Acesso em: 17 abr. 2004.

SANTAREM SEGUNDO, J. E. *Recursos tecno-metodológicos para a descrição e recuperação de informações na web*. 2004, 157 f. Dissertação (Mestrado em Ciência da Informação) – Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília, 2004.

SANTOS, P. de M. L. dos. *O ponto de inflexão Otlet: uma visão sobre as origens da documentação e o processo de construção do Princípio Monográfico*. 2006, 146f. Dissertação (Mestrado em Ciência da Informação) – Escola de Comunicação e Artes, Universidade de São Paulo, São Paulo, 2006. Disponível em: http://poseca.incubadora.fapesp.br/portal/bdtd/2006/2006-me-santos_paola.pdf. Acesso em: 08 de jan. 2007.

SARACEVIC, T. Ciência da informação: origem, evolução e relações. *Perspectiva Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.

SOUZA, R. R.; ALVARENGA, L. A Web Semântica e suas contribuições para a ciência da informação. *Ciência da Informação*, Brasília, v. 33, n. 1, p. 132-141, jan./abr. 2004.