

O campo da ciência da informação: contribuições, desafios e perspectivas

Januário Albino Nhacuongue

Doutor em Ciência da Informação pela Universidade Estadual Paulista "Júlio de Mesquita Filho" - UNESP, Campus de Marília. Mestre em Ciência da Informação pela UNESPLicenciado em Ciências Policiais pela Academia de Ciências Policiais de Moçambique - ACIPOL

Edberto Ferneda

Pós-doutorado pela Universidade Federal da Paraíba .Mestre em Informática pela Universidade Federal da Paraíba Doutor em Ciências da Comunicação (Ciência da Informação) pela Universidade de São Paulo. Graduado em Processamento de Dados pela antiga Fundação Educacional de Bauru

<http://dx.doi.org/10.1590/1981-5344/1932>

A Ciência da Informação é um campo que ainda denota traços atinentes à sua estruturação e solidificação, por isso, ainda sofre influências externas devido ao respectivo fraco grau de autonomia. A sua gênese é atribuída à explosão informacional que caracterizou o período pós - segunda Guerra Mundial e às conferências subsequentes que visavam analisar a implementação de máquinas na transferência da informação. Parte do seu contributo social é centrada na recuperação da informação mediada pela tecnologia que garante a despersonalização do conhecimento para indivíduos, grupos e sociedade, através da informação como conhecimento em ação. A área desenvolveu-se sobre um enfoque tecnológico de caráter interdisciplinar que atualmente cria as suas limitações. A informação registrada que constitui o seu objeto pela possibilidade de manipulação, também se torna uma limitação, daí a necessidade de adoção da técnica de mineração de dados para a descoberta do conhecimento em grandes volumes de dados e ampliação do universo da recuperação da informação.

Palavras – chave: *Ciência da Informação; Tecnologia; Campo Científico; Mineração de dados; Conhecimento.*

The field of information science: contributions, challenges and perspectives

Information Science is a field that still denotes traits peculiar to its structuring and solidification. Therefore still suffers external influences due to its low degree of autonomy. Its genesis is attributed to the information explosion that characterized the post - World War II and the subsequent conferences aimed at reviewing the implementation of machines in the information transfer process. Part of its social contribution is centered in information retrieval mediated by technology that ensures depersonalization of knowledge to individuals, groups and society through information and knowledge into action. The area has developed a focus on interdisciplinary technology that actually creates its limitations. The recorded information that constitutes its object by the possibility of manipulation, it also becomes a limitation, hence the need to adopt the technique of data mining to knowledge discovery in large data volumes and expansion of the universe of information retrieval.

Keywords: Information Science. Technology. Scientific field. Data Mining. Knowledge.

Recebido em 26.11.2013 Aceito em 12.05.2015

1 Introdução

O advento da ciência moderna que é atribuída a Galileu, ganhou maior repercussão científica e social através do método científico que, por conseguinte, constitui a sua lógica. Com o método científico, a validade do conhecimento passou a se basear na possibilidade de observação, experimentação e explicação por meio de teorias ou processos comuns. Com o método, mais pessoas são chamadas a intervir sobre o conhecimento alheio, quer agregando-lhe valor substancial, quer fazendo interseções que tendem a aceitá-lo ou contrastá-lo.

A tecnologia desempenha um papel preponderante na universalização da informação, através de ambientes que permitem a interação contínua no compartilhamento. Todavia, a priori, a sua eficácia está condicionada à eficiência do seu processo, dos seus métodos e do seu caráter, ou seja, em primeiro lugar a tecnologia deve ser autossuficiente e

universalmente permeabilizadora, para depois contemplar aspectos sobre a individualidade dos sujeitos.

A Ciência da informação teve uma gênese principalmente impulsionada pelo movimento acelerado das tecnologias de informação e comunicação. O seu desafio é notabilizado tanto por meio de abordagens sobre a natureza, manifestações e efeitos da informação e conhecimento, como pela busca de soluções tecnológicas que garantam a comunicação e uso. Por isso que o objeto (informação) inclui todo o ciclo (produção, organização, armazenamento, representação, disseminação, recuperação, acesso e o uso) da informação.

A natureza interdisciplinar da Ciência da Informação com enfoque tecnológico tem sido égide de outros aspectos complexos e limitativos sobre a sua atuação, principalmente pelo próprio caráter reducionista da tecnologia e pela fraca abrangência das suas ferramentas na transferência e nos fluxos informacionais.

Numa abordagem qualitativa, este trabalho elucida os diversos aspectos que caracterizaram a gênese da Ciência da Informação como campo científico e os seus efeitos na atualidade, propondo soluções para potencializar a transformação social pela informação.

2 O campo científico

A Ciência da Informação pode ser abordada com muita propriedade sob o ponto de vista estrutural e funcional. Porém, sob o prisma epistemológico, tal tarefa se torna complexa por suscitar uma análise profunda dos aspectos nos quais a área se encontra enraizada e que lhe permitem uma discussão concomitante com outros campos, sobre a natureza e a validade do conhecimento humano.

Para elucidar a essência da Ciência da Informação no seio de outras áreas de conhecimento, este trabalho parte da noção de campo proposta por Pierre Bordieu.

Bordieu (2004) fez uma reflexão sobre a lógica do mundo científico, apontando o campo científico como um microcosmo ou espaço de atuação entre agentes (indivíduos e instituições) que produzem, reproduzem e difundem a ciência. Este espaço é caracterizado por leis próprias, através das quais resiste a pressões externas e que lhe conferem o respectivo grau de autonomia, isto é, "quanto mais os campos científicos são autônomos, mais eles escapam às leis sociais externas". Por outras palavras, campos autônomos apenas sofrem críticas e contrastes através de argumentos e demonstrações; enquanto que os campos sem normas e estrutura rígidas são facilmente manipulados por meio de intervenções não científicas (construções sociais). Por isso, quando os problemas políticos são notórios no seio das disciplinas de um campo isso se torna um indício da fraca autonomia desse campo no todo.

Uma das características do campo, segundo Bordieu (2004), é a existência de relações de forças consubstanciadas através das relações objetivas entre os seus agentes. No caso do campo científico, a pesquisa, a produção literária e o lugar da difusão dessa produção são

consequências dos interesses e da posição tomada pelos seus agentes, ou seja, os agentes “fazem os fatos científicos”, e conseqüentemente, o campo científico.

Cada campo tem a sua forma específica de capital, o capital simbólico que se constitui a partir de “atos de conhecimento e reconhecimento”, por meio de citações, prêmios, etc. Deste modo, Bordieu (2004) destaca duas espécies de capital científico, por um lado o temporal ou político, institucional e institucionalizado, ligado à ocupação de posições estratégicas nas instituições científicas e aos meios de produção e reprodução. Por outro, o poder específico ou de prestígio pessoal resultante do reconhecimento do agente pelas suas qualidades científicas. Enquanto o capital institucional se fortalece através de estratégias políticas de participação em reuniões, eventos, bancas, comissões, entre outras, ele se torna legítima e inquestionável; em contraste, o capital pessoal depende da contribuição reconhecida do agente em prol da ciência, por meio de invenções ou pesquisas e publicações em órgãos de renome, e por isso, se torna passível de contraste e crítica.

O problema de autonomia e das relações de forças existentes no campo científico é resumido por Saracevic (1996) como sendo o problema da ecologia informacional. A ecologia enquanto espaço de coexistência entre produtores, editoras, revistas, instituições de fomento de pesquisas, etc., é caracterizada por conflitos de interesse e que indubitavelmente afetam o modo de produção e divulgação da ciência nas diferentes áreas de conhecimento, no geral e na Ciência da Informação, em particular.

A Ciência da informação é uma ciência nova e em fase de estruturação. Este fato é testemunhado por Smit *et al.* (2004), que afirmam que além da falta de consenso no que tange ao seu objeto e delimitação, o próprio termo “Ciência da Informação” ainda se apresenta como insignificante cujo significado está em construção. Como tal, pela falta da estrutura básica e conceitual, ainda sofre influências externas porque o capital científico institucional pesa mais sobre os poucos pesquisadores renomados da área. Ademais, especificamente no Brasil, há uma tendência generalizada de influência e controle da academia através das instituições de fomento de pesquisa, que além de classificar os programas de pós-graduação, avaliam o desempenho anual de cada programa, por meio de vários requisitos, dentre eles a publicação nos periódicos por eles também classificados. Alia-se à política de controle acadêmica, a editoria da maior parte desses periódicos por pesquisadores consagrados na área que, tanto de forma implícita ou explícita, contribuem para a temática das publicações pelo critério de aceitabilidade ao escopo do periódico. Por isso, a Ciência da Informação e outros campos do conhecimento ainda se mostram desfalecidos do real contributo ou dos respectivos “usos sociais”, na medida em que a maioria das temáticas de pesquisa encontra-se “politizada” pelo capital científico institucionalizado e monopolista dos agentes que ocupam as posições estratégicas.

3 A Ciência da Informação

A Ciência da Informação, conforme anteriormente se referiu, é um campo científico em construção. Este fato é notório nas divergências que ainda predominam sobre a sua gênese e conceituação da maioria das suas terminologias. Em relação à sua gênese, alguns pesquisadores como Rayward (1994) e Barreto (2008) são apologistas do entendimento que vislumbra os traços básicos da área muito antes da explosão informacional que caracterizou o período Pós-Segunda Guerra Mundial. Em contrapartida, Capurro e Hjørland (2003), Wersig (1993) e Saracevic (1996), associam a origem da Ciência da Informação à busca de soluções para resolver o problema informacional no período Pós-Guerra.

Numa perspectiva histórica, Barreto (2008, p. 2) considera que a gênese da Ciência da Informação subjaz da inovação tecnológica e que “o fluxo de informação e sua distribuição ampliada e equitativa” sempre motivaram o homem, desde a fase das pinturas rupestres, passando pela invenção da escrita até a Internet. Neste contexto, embora de forma indireta, percebe-se que o autor relaciona a busca de soluções sobre a produção, organização e disseminação da informação, ao papel das redes de saber universal, como a Academia de Lince de 1603 e a *Royal Society* de 1660. Além das redes de saber universal e das redes de conhecimento distributivo (enciclopédias), é na rede universal de conhecimento de Otlet e La Fontaine que o autor advoga o marco importante da Ciência da Informação:

os determinantes colocados anteriormente permitem refletir com mais liberdade a questão da Ciência da Informação em um desenrolar histórico descritivo, que tem somente a validade no contexto do desenvolvimento histórico da informação e conhecimento. Permitem ainda verificar que o ideal do acesso ao conhecimento livre e para todos não surgiu com a Internet (BARRETO, 2008, p. 3).

De fato, muitas questões que norteiam as diretrizes de atuação da área, como o livre acesso, a rede universal de conhecimento, os usos e usuários, entre outras, foram incorporadas ou retomadas a partir dos estudos de Paul Otlet. Esta é a percepção que se obtém de Rayward (1994), ao considerar que foi Otlet que desenvolveu o sistema complexo de organização para integrar bases de dados bibliográficos, imagéticos e textuais que atualmente funcionam como hipertexto. Por isso, o Tratado de Documentação de Otlet de 1934 deveria ser a primeira obra na Ciência da Informação, na medida em que condiz com muitos aspectos do campo, como a busca multifacetada e integrada do conhecimento humano por nós associativos.

Para Barreto (2008), Vannevar Bush (1945) pode ser considerado o pioneiro da Ciência da Informação e 1945 sua data fundadora pela publicação do seu artigo, principalmente pelas conferências subsequentes para a discussão das suas ideias, como a Conferência da *Royal Society* de

1948 e que culminaram com a criação da área como um novo campo científico para lidar com a nova situação informacional de emprego de máquinas para simular o processo racional humano de busca de informações por associação conceitual, terminológico e linguístico.

A maioria dos autores aponta a explosão informacional, essencialmente marcado pela quantidade de relatórios que documentavam a Segunda Guerra Mundial e as conferências para debater a solução tecnológica proposta por Vannevar Bush como o marco inicial da Ciência da Informação. Segundo Capurro e Hjørland (2003), a Ciência da Informação nasceu nos anos 50, aliada ao desenvolvimento e disseminação de uso de computadores, no final da Segunda Guerra Mundial.

O entendimento de Wersig (1993) coaduna com a percepção de Barreto (2008), sobre a origem da Ciência da Informação enleada à inovação tecnológica, porém Wersig (1993) acrescenta que o seu fundamento se encontra na explosão informacional pós-guerra. Neste sentido, o autor aponta que ao contrário das ciências clássicas cuja gênese está enraizada na busca do entendimento completo de como o mundo funciona, a Ciência da Informação nasce na necessidade de desenvolver estratégias para resolver problemas causados por ciências e tecnologias do clássico, impondo mudanças significativas no conhecimento, como visão técnico-sistema para visão usuário/humano.

Ao contextualizar a origem da Ciência da Informação a partir do problema da informação, Wersig (1993) aponta que a área nasceu no Século XX com a documentação que procurava solucionar o problema do "dilúvio da literatura", passando pelo foco na recuperação da informação e culminou com aspectos complexos da tecnologia, tornando-se ciência. Assim, a área nasce na mudança do papel do conhecimento para indivíduos, organizações e culturas, influenciada em grande parte, pelo advento da tecnologia. Um dos efeitos desta mudança foi a despersonalização do conhecimento, através da tecnologia da comunicação (imprensa) que permitiu que o conhecimento oral fosse traduzido em suportes e disseminado pela multiplicidade de usuários.

Saracevic (1996) traçou o perfil da Ciência da Informação a partir de três características: a interdisciplinaridade na sua natureza, a sua ligação com a tecnologia e o seu papel ativo na evolução da sociedade da informação. Assim, o autor partilha a corrente que atribui à gênese do campo o movimento científico que procurava solucionar o problema informacional no período pós-guerra e considera que o artigo *as we my think*, publicado por Vannevar Bush (1945), caracterizou o embrião da Ciência da Informação. Esta afirmação é embasada na proposta tecnológica de associação automatizada de ideias na recuperação para colmatar o problema de excesso da informação e pelos diversos esforços conjugados posteriormente em prol da importância estratégica da informação para indivíduos, grupos e organizações.

Para Saracevic (1996) a recuperação da informação caracterizou a essência da Ciência da Informação na sua origem e desenvolveu-se sob três vertentes: descrição intelectual da informação, especificação intelectual da busca e especificidade sobre sistemas e técnicas a empregar. Deste modo, a área emergiu como campo científico nos anos 60 e os seus problemas estão sintetizados tanto na

natureza, manifestações e efeitos da informação e conhecimento, como nos processos da comunicação e uso da informação.

A Ciência da Informação nasce num panorama que prevê, por um lado, a geração e a organização da informação e, por outro, a transferência da informação mediada pela tecnologia para a sua preservação e recuperação:

a Ciência da Informação é a disciplina que investiga as propriedades e o comportamento informacional, as forças que governam os fluxos de informação, e os significados do processamento da informação, visando à acessibilidade e a usabilidade ótima. [...] está preocupada com o corpo de conhecimentos relacionados à origem, coleção, organização, armazenamento, recuperação, interpretação, transmissão, transformação. [...] uso de códigos para a transmissão eficiente da mensagem, bem como o estudo do processamento e de técnicas aplicadas aos computadores e seus sistemas de programação (BORKO, 1968, p. 1).

Esta conceituação pode ser resumida através das cinco áreas de atuação de que Ingwersen (1992) se refere, nas quais se concentram as suas disciplinas: transferência da informação (fluxo de informação); informação desejada (relação entre a produção da informação e as necessidades entre indivíduos, grupos ou sociedade); efetividade dos sistemas (tecnologia para tornar eficiente a qualidade da informação); a relação entre a informação e o gerador (geração do conhecimento e a sua representação) e a relação entre a informação e o usuário (relevância, valor e uso da informação).

Embora as controvérsias em relação à origem da Ciência da Informação prevaleçam, parece ser pacífico afirmar que o campo se institucionalizou como ciência no período pós-segunda guerra. Ademais, muitas das discussões atinentes a sua revitalização foram desencadeadas nesse período. Este fato é testemunhado por Ingwersen (1992) que cita a *Conference on Conceptions of Library and Information Science (CoLIS)* de 1977 como sendo a pioneira em que se evidenciou um aprofundamento sobre questões básicas, como transferência da informação e processos de comunicação humana, mediados pela tecnologia.

A Ciência da Informação é definida a partir dos problemas que se predispõe a resolver e a sua estrutura está baseada no vínculo permanente entrelaçado entre a pesquisa empírica e a prática profissional. Esta característica justifica-se na origem do campo, enquanto proposta inovadora na ciência de uma área técnica específica em resposta ao problema do excesso da informação. Por isso a sua relação trans-inter-multidisciplinar com a Biblioteconomia, Ciência da Computação, Documentação, Inteligência Artificial é justificável e necessária, na medida em que o problema do excesso da informação se torna cada vez mais complexo com a *Web*.

4 A contribuição para o conhecimento

A Ciência da Informação desempenha um papel social preponderante, não apenas pelo foco na dimensão humana na relação com a tecnologia, como também no processo de transformação pela informação, isto é, naquilo que Wersig (1993) denomina "informação como conhecimento em ação". Equivale dizer que atualmente o campo se empenha em soluções para lidar com a despersonalização do conhecimento e sua fragmentação e no desenvolvimento de outras formas adequadas de racionalização que estejam abertas a todos os tipos de conhecimento. É sobre esta característica que reside a peculiaridade da dimensão social da área reconhecida por Saracevic (1996), que é de índole social global e transcendental do aspecto humano.

Engelbert (2003) alertava sobre a necessidade do contributo social da ciência e tecnologia através da inovação descontínua e computação interativa, de modo a tornar colaborativa a construção do conhecimento. É por meio deste tipo de inovação que a Ciência da informação quebrou os paradigmas que caracterizavam outros campos científicos e se desenvolveu como uma ciência de natureza essencialmente técnica para solucionar os problemas informacionais da sociedade. Por isso que o campo reforça a mudança ideológica pós-modernista do foco nas pessoas e não nos ideais abstratos teóricos e doutrinários do Homem.

O contributo da Ciência da Informação é visível nas esferas informacionais de caráter científico, político, administrativo, econômico, social, entre outras. Mas não se pode exacerbar ao extremo sobre isso. Morin (2003) apela para o abandono das correntes radicais que tendem a analisar o único lado do problema. Conforme o autor acrescenta, não existe o todo sem as partes, da mesma forma em que não existem as partes sem o todo, ou seja, é preciso um pensamento complexo que analise o assunto e as suas respectivas facetas. Por isso, embora aprofundado no ponto sobre o objeto de pesquisa da Ciência da Informação, aqui se discutem alguns problemas da área para o conhecimento. De referir que a maioria dos problemas descritos não é somente da Ciência da Informação, mas da ciência no geral. Porém, estes problemas são solidariamente notórios nesta pela sua natureza essencialmente enleada à tecnologia.

Uma das questões levantadas por Saracevic (1996) se prende com o modo como o campo da Ciência da Informação se desenvolve, construindo a base dos seus problemas a partir do enfoque tecnológico, em vez do enfoque humano. O autor questiona esta abordagem da área que subsidia a implementação dos sistemas de acesso à informação pelos critérios de eficiência e relevância baseados nas aplicações tecnológicas, na medida em que tais critérios são estritamente humanos. De fato, não se pode pensar na neutralidade da tecnologia, do mesmo modo em que não existe neutralidade na mediação da informação para o processo interativo de comunicação humana. Feenberg (2010, p. 206) afirma que a "tecnologia continuará a afetar cada vez mais a vida social e cada vez menos permanecerá livre da sua influência na constituição de uma diferença cultural". Esta percepção é baseada no determinismo tecnológico Heideggeriano, caracterizado por uma transformação do tecido social e

consequente instrumentalização do homem em função da técnica, a tal ponto que não reste alternativa para quem a ela se abstenha. Com a sofisticação da tecnologia, caminha-se para um futuro em que a informação estará totalmente digitalizada e conseqüentemente o suporte será o banco de dados, restando para o usuário do papel apenas a conformidade.

Um dos problemas apontados por Wersig (1993) e que ainda são objetos de análise na Ciência da Informação se prende com a fonte e apropriação da informação. Com a tecnologia, a fonte de informação se torna cada vez menos evidente. Esta situação é comum na *Web* e acaba por ofuscar os ganhos da referida despersonalização do conhecimento, já que informações importantes são preteridas pela falta da clareza da respectiva fonte ou informações não confiáveis são usadas para assuntos complexos e sensíveis. De igual modo, na mediação da informação, algumas organizações se apropriam dessa informação como se delas se tratasse, prejudicando a visibilidade dos reais autores, o acesso e uso para outras finalidades.

O segundo efeito da mudança do papel do conhecimento apontado por Wersig (1993) está relacionado à credibilidade do conhecimento, através da tecnologia que permite a observação. Se no passado o conhecimento produzido poderia ser testado por muitos pesquisadores, atualmente, com a sofisticação das teorias, metodologias e técnicas de coleta e processamento de dados, o conhecimento produzido fica restrito a certos grupos ou comunidades.

Se mais conhecimento se torna despersonalizado, por outro lado, mais conhecimento tem de ser acreditado [...]. A situação se tornará mais complicada com as novas tecnologias [...]. Por isso, cada vez mais temos que ter cuidado com os dados de observação em dois aspectos: em primeiro lugar temos que aceitar a tecnologia que originou os dados, e depois levar em conta o que poderia ter acontecido com os dados brutos em processo de transformação. Para aceitar o conhecimento temos que ser muito críticos em relação às tecnologias de coleta e manipulação (WERSIG, 1993, p. 232).

Por isso, atualmente a Ciência da Informação procura soluções para esse problema através da política de acesso a dados abertos, de modo a subsidiar pesquisas afins, envolvendo outras ferramentas de manipulação e observação. Além de garantir a prova das conclusões tiradas pelo autor dos dados, o processo permite a reutilização dos dados para outras finalidades, daí a importância da descrição detalhada de todo o processo envolvido na coleta desses dados.

O terceiro efeito da mudança do papel do conhecimento segundo Wersig (1993) é a fragmentação do conhecimento por meio da tecnologia de apresentação. O autor aponta três razões dessa fragmentação: volume excessivo; autonomização das áreas de conhecimento, através de padrões próprios de cada campo ou disciplina e sistemas finais, enquanto

ideologias de pensamentos sobre o mundo. Estes fatores culminam com a “diversificação de tecnologias de apresentação de conhecimento”, levando as pessoas a um conhecimento fragmentado de acordo com as suas crenças, culturas e áreas de formação ou atuação. Neste contexto, verifica-se uma concorrência desleal no conhecimento desfalecido na “maquilhagem” do fator potencializador e globalizante da tecnologia. Ora, os países do terceiro mundo que ainda lutam para suprir os elevados índices de analfabetismo com sistemas deficitários de educação são inseridos na falsa sociedade da informação, onde os seus agentes (usuários) quase pouco ou nada produzem e a sua visibilidade no sistema é imperceptível para os países com grandes economias.

Aliado à situação da fragmentação está o fator língua. A maioria das pesquisas técnicas com elevado grau de aprofundamento são dos países mais ricos, já que detêm melhores estruturas de pesquisa e desenvolvimento e, por isso, são disponibilizadas em inglês. Além da obrigatoriedade do aprendizado desta língua para o compartilhamento do conhecimento dos poucos conteúdos disponíveis pela parte dos países mais pobres, constata-se uma fraca ou recíproca nula, na medida em que a produção científica destes quase não suscita nenhum interesse pela parte dos mais ricos. Em parte, isto acontece porque a informação e o conhecimento estão entrelaçados com os meios de produção e, por isso, tornaram-se instrumentos capitalistas de dominação e controle.

O quarto efeito da mudança do conhecimento segundo Wersig (1993) é a racionalização do conhecimento, através da tecnologia da informação. Uma das razões para essa racionalização se prende com o aumento de tecnologias para a disseminação e acesso à informação e das potencialidades de uso do conhecimento para a solução dos problemas da sociedade. O problema é que nem todos os usuários com necessidades de informação têm acesso à tecnologia e, por isso, eles permanecem excluídos circunstancialmente, daí a aparente utopia sobre o real alcance da sociedade da informação e do conhecimento.

Parece que a Ciência da Informação enveredou pela alternativa crítica levantada por Saracevic (1996), que consiste na facilidade de ensinar e ajustar os humanos à máquina do que criar aplicações tecnológicas efetivas e que correspondam aos modelos mentais humanos. A área deve buscar com premência um equilíbrio na avaliação da relação homem – tecnologia, para não ser considerada como um campo ideal para a indústria da informação, mas um fracasso para a sociedade. Para isso, pequenas ações intervencionistas de caráter antropológico, voltadas para pequenas comunidades serão salutárias para a dimensão social.

Os estudos de usuários que, segundo Figueiredo (1994, p. 7), são “investigações que se fazem para saber o que os indivíduos precisam em matéria de informação, ou então, para saber se as necessidades de informação estão sendo satisfeitas de maneira adequada”, devem potencializar o comportamento e as necessidades dos usuários do que os centros de informação. Este é outro entrave que o campo enfrenta na atualidade da *Web*, tanto pela dificuldade de identificar os usuários de

determinados serviços de informação, como pela impossibilidade da sua estratificação.

5 O problema do objeto de pesquisa

O caráter controverso em relação ao objeto de pesquisa da Ciência da Informação e a respectiva conceituação transparece o fulcro do problema para a área. Smit *et al.* (2004) afirmam que tal se deve ao fato do campo assumir o seu desenvolvimento a partir de uma estrutura interdisciplinar de índole funcional, isto é, a maioria dos conceitos empregues na área foi apropriada a partir de outras áreas em razão dos propósitos que procuravam satisfazer na Ciência da Informação. De fato, a Ciência da Informação nasce numa vertente essencialmente tecnicista do que empírica e por isso a sua contribuição é notória nas pesquisas que visam solucionar problemas.

A falta da base estrutural e conceitual da Ciência da Informação e a relutância em razão do objeto "informação", tanto pela falta da clareza sobre o seu conceito, como por parecer assunto tratado por outras áreas, não nos deve alarmar, pois segundo Wersig (1993), trata-se de uma ciência pós-moderna e como tal, não precisa de um método específico porque lida com uma diversidade de problemas, causados pelas ciências clássicas e pela tecnologia, nos quais não caberia apenas um único método. A Ciência da Informação não deve ter uma estrutura à semelhança dos campos tradicionais.

Mesmo caminhando no panorama nebuloso de incertezas e controvérsias, assume-se tendencialmente "a informação" como o objeto de estudo da Ciência da Informação. O significado do termo informação é analisado na relação com o conhecimento por Popper (1972 *apud* BROOKES, 1980), através de três mundos: o mundo 1 correspondente ao mundo físico, isto é, da informação como a essência do universo ou da terra; o mundo 2 correspondente ao conhecimento humano subjetivo, ou seja, da informação apenas no estado mental do sujeito, segundo as suas capacidades perceptivas individuais criadas por Deus ou outra força cósmica e por último o mundo 3 ou mundo do conhecimento objetivo, da informação como produto dos mundos 1 e 2, isto é, das reflexões, experiências e observações sobre o universo, cujo conteúdo se apresenta registrado através de linguagens, artes, ciências, tecnologias, cultura ou outras manifestações. Enquanto outros campos tratam especificamente dos mundos 1 e 2, é no mundo 3 ou no conhecimento objetivo que se centra o objeto da Ciência da Informação.

A ambiguidade do termo informação também é analisada por Buckland (1991) com um aperfeiçoamento crítico considerável, pela especificação dos diferentes usos da informação. A tripartição do autor se baseia nos três momentos da informação, sendo o primeiro dos processos - informação como processo, isto é, o ato de informar, comunicar ou falar algo. Em segundo, a informação como conhecimento, resultante da apreensão ou percepção de algo durante os processos anteriores. Por último, a informação como coisa, ou seja, exteriorização do que foi

percebido ou apreendido em objetos informativos que contêm a qualidade de conhecimento comunicado, comunicação, informação ou algo informativo. O primeiro momento da informação corresponde ao contato intrínseco entre o sujeito e objeto, o segundo, a uma transformação do sujeito pela aquisição de algo intangível na sua mente e o terceiro, à exteriorização do intangível em algum suporte tangível.

Como se pode depreender, a Ciência da Informação trabalha com a informação registrada ou conhecimento objetivo. Esta característica intencional justifica-se pela sua funcionalidade, isto é, da possibilidade de manipulação mediada por sistemas tecnológicos para a interação com o usuário do processo comunicativo.

Ao enaltecer o progresso da Ciência da Informação, Ingwersen (1992) reconhece que ao considerar o usuário da informação o público humano no lugar da máquina, a área expandiu claramente o seu escopo na sociedade, daí a necessidade do aprofundamento das questões sobre o uso e transformação da informação em conhecimento nos níveis individual e social. Ora, ao mesmo tempo em que o campo se torna evidente pelo seu papel social, a limitação dos processos e técnicas embasados na informação registrada se revela como o seu maior problema na atualidade. Equivale por outras palavras dizer que com a *Web* muitas das informações que a área tem de lidar ficam perdidas pela impossibilidade de manipulá-las. Esta limitação é evidenciada essencialmente pela intangibilidade do suporte na qual as informações se encontram representadas.

A limitação da Ciência da Informação em relação ao suporte suscita cada vez mais a interdisciplinaridade que caracteriza a natureza da área. Neste sentido, este trabalho sugere a adoção de tecnologias emergentes e que permitem sintetizar as informações para futuros domínios ou perspectivas específicas.

6 Tecnologias para o acesso à informação

Conforme anteriormente se referiu, a Ciência da Informação evidencia-se notavelmente no acesso à informação, mormente através da recuperação da informação. Porém, os atuais processos, métodos e técnicas adotados pela área revelam uma grande limitação para lidar com o crescente volume de dados na *Web*.

O conceito de *Big Data* surge como uma das alternativas para o processamento de grandes e complexas coleções de dados, face à incapacidade das tradicionais ferramentas de gerenciamento de bancos de dados. O processo visa aperfeiçoar a captura, a curadoria, o armazenamento, a busca, o compartilhamento, a transferência, a análise e a visualização dos dados, principalmente para ambientes corporativos. *Big Data* se desenvolve sob o lema de "3V's", isto é, alta velocidade de coleta e processamento, alto volume de dados e alta variedade de ativos de informação que requerem novas formas de processamento para fundamentar decisões ou permitir a descoberta do conhecimento. Ultimamente, a veracidade e o valor têm sido acrescidos ao processo.

Big Data é usada em *business intelligence*, através de estatísticas que permitem fazer medições e detectar tendências ou para fazer associações, regressões, etc. que demonstrem relações ou previsões de resultados. Esta tecnologia é de suma importância pela capacidade de processamento de grandes volumes de dados e inferências em curto prazo. Contudo, para a Ciência da Informação, o processo ainda se mostra complexo, não só por suscitar grandes plataformas tecnológicas com softwares sofisticados (algoritmos genéticos, processamento em linguagem natural, redes neurais, processamento de sinais, etc.), como também por requerer conhecimentos específicos de quadros especializados. Além disso, o processo tem sido severamente criticado por alguns analistas, pois as inferências estatísticas são feitas com base na diversidade de informações colhidas em diferentes períodos, locais e situações, descontextualizadas da origem e complexidade envolvida na sua coleta. Outra crítica importante se prende com a invasão constante da privacidade, na medida em que dados pessoais são coletados, integrados e armazenados em estruturas únicas para apoiar interesses singulares de empresas, instituições, laboratórios, entre outros.

A mineração de dados, principalmente a mineração de dados da *Web* ou *Web mining* é outra tecnologia emergente que visa potencializar o acesso à informação.

A mineração de dados, à semelhança da Ciência da Informação, é uma área que surgiu com o problema do excesso da informação, neste caso, no desafio de responder ao problema do elevado volume de dados armazenados em bancos, através de identificação de padrões e de regras significativas a esses dados. Tal desafio surgiu para resolver o defasamento das técnicas tradicionais de análise de dados, face à nova característica da complexidade dos dados.

Tan, Steinbach e Kumar (2009, p. 3) consideram que “a mineração de dados é o processo de descoberta automática de informações úteis em grandes depósitos de dados”. A partir desta conceituação, podem-se destacar quatro elementos essenciais: processo, descoberta automática, informações úteis e grandes depósitos de dados. Como processo, significa que a mineração de dados é um método ou sistema com regras específicas. A descoberta automática implica a obtenção de resultados por meios puramente algorítmicos, sendo por isso, em algumas vezes, resultados imprevisíveis à cogitação humana. O elemento informações úteis significa que os resultados obtidos devem ser proveitosos para a tomada de decisões. Por último, o elemento depósito de dados implica, a priori, a existência de um sistema gerenciador de banco de dados para o armazenamento, indexação e processamento de consultas.

O objetivo final da mineração de dados é o processo *Knowledge Discovery in Databases* – descoberta do conhecimento em banco de dados. O processo pressupõe numa primeira fase a entrada e armazenamento de dados. Em seguida, o pré-processamento para transformar os dados brutos em formato apropriado para a análise por meio das seguintes tarefas: fusão de dados de múltiplas fontes, limpeza de dados para a remoção de ruídos, observações duplicadas, seleção de

registros e características relevantes para a mineração. A fase da mineração em si corresponde à identificação de padrões de dados e integração com ferramentas de apoio a decisão. As tarefas da mineração consistem em prever valores de atributos dependentes a partir de valores de outros atributos independentes ou derivar padrões pelos quais os dados se relacionam. Para tal, constrói-se o modelo de previsão com base na classificação ou regressão das variáveis ou se recorre a associações ou agrupamentos das características dos dados (TAN; STEINBACH; KUMAR, 2009).

A mineração de dados também pode ser usada para melhorar os sistemas de recuperação da informação, convertendo numerosos dados brutos em informações úteis e agregados de valor que, a priori, não fariam nenhum sentido para o usuário. Algumas informações disponíveis na *Web* não podem ser manipuladas pelos sistemas de recuperação da informação, pela dificuldade do acesso aos respectivos repositórios ou suportes. Neste caso, a mineração de dados é importante para a Ciência da Informação, pois amplia o âmbito do seu objeto, pela possibilidade de recuperação de informações ricas em metadados que podem ser preservadas e reutilizadas em diferentes unidades ou contextos informacionais. Por exemplo, diariamente informações importantes são veiculadas através do *Twitter*, do *Facebook*, do *LinkedIn* e de outras mídias sociais e, indubitavelmente, ficam fora do escopo da Ciência da Informação. Com as ferramentas de mineração de dados da *Web* social é possível saber os assuntos comentados no *Twitter*, identificando as coordenadas espaciais dos respectivos autores, calcular as *trends*, calcular a diversidade léxica para *tweets*, fazer pesquisas com expressões regulares para encontrar *retweets*, criar e analisar grafos sobre *retweets*, analisar a conexidade na rede com detalhes de visualização do *RGraph* ou *WP-Cumuls*, entre outras ações. Todas estas informações podem ser armazenadas em repositórios próprios para futuras manipulações e finalidades da área. A mineração de dados é uma componente que pode ser agregada à recuperação da informação para ampliar o universo informacional, mediante a prospecção de outras possibilidades do conhecimento.

7 Considerações finais

A Ciência da Informação precisa buscar com urgência a sua autonomia para contribuir com certo grau de liberdade com respostas às diversas necessidades informacionais da sociedade, por meio de pesquisas empíricas, teóricas e das técnicas e dos processos evidenciados pela atuação dos diversos profissionais.

Embora a visibilidade da área seja notória em vários aspectos que envolvem a linguagem, a comunicação e a epistemologia, é na representação e recuperação da informação mediada pela tecnologia em que tais abordagens se justificam. Por isso, sendo a tecnologia “de per si” um fator limitativo, a materialidade da informação se torna condicional e o conhecimento em ação uma probabilidade no infinito das possibilidades

informativos. Deste modo, o foco da área baseado na dimensão humana contrasta a atual abordagem centrada na tecnologia.

A abordagem da Ciência da Informação recai num dilema, pois ao mesmo tempo em que se levantam vários questionamentos sobre a constante simbiose com a tecnologia, a dinâmica atual torna complexa e impensável a tarefa de criar um campo destituído da componente tecnológica. Tal complexidade foi herdada da própria gênese do campo, enquanto proposta científica para resolver o problema do excesso da informação, num processo baseado no uso de ferramentas tecnológicas. Por isso que a maioria dos pesquisadores pioneiros consagrados na Ciência da Informação e principalmente na recuperação da informação era oriunda de áreas, como Ciência da Computação, Engenharia, Matemática e Física. Entre esses pesquisadores destacam-se, segundo Chu (2007): Hans P. Luhn que era formado em engenharia e se evidenciou com a indexação automática e resumos e disseminação seletiva da informação; Calvin N. Mooers, formado em Matemática e Física e cunhou o termo *information retrieval* em 1950; Karen S. Jones, formada em história, filosofia e computação se destacou com o uso da estatística para a frequência das palavras-chave e frequência do documento inverso na busca e processamento em linguagem natural.

A mineração de dados poderá potencializar o acesso à informação, incluindo proporções que atualmente não são atingidas pela recuperação da informação, e que configuram a dimensão social das necessidades específicas do usuário.

Referências

- BARRETO, A. A. Uma quase história da ciência da informação. *Datagramazero - Revista de Ciência da Informação*, Rio de Janeiro, v. 9, n. 2, abr. 2008. Disponível em: <http://www.dgz.org.br/abr08/Art_01.htm>. Acesso em: 17 abr. 2013
- BORKO, H. Information science: what is it? *American Documentation*, v. 19, n. 1, p. 3-5, jan. 1968.
- BOURDIEU, P. *Os usos sociais da ciência: por uma sociologia clínica do campo científico*. São Paulo: Editora UNESP, 2004.
- BROOKES, B. C. The foundations of Information Science: part I - philosophical aspects. *Journal of Information Science*, v. 2, n. 3-4, p. 125-133, 1980. Disponível em: <<http://jis.sagepub.com/content/2/3-4/125.full.pdf+html>> Acesso em: 10 maio 2009.
- BUCKLAND, M. K. Information as thing. *Journal of the American Society for Information Science (JASIS)*, v. 45, n. 5, p. 351-360, 1991.
- BUSH, V. As we may think. *Atlantic Monthly*, v. 176, n. 1, 1945. Disponível em: <<http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/>>. Acesso em: 3 mar. 2010.

CAPURRO, R.; HJORLAND, B. O conceito de informação. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 12, n. 1, jan./abr. 2007. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54/47>> Acesso em: 13 jun. 2013.

CHU, H. *Information representation and retrieval in the digital age*. 3ª Tiragem. New Jersey: Asist&T, 2007.

ENGELBART, D. C. *Improving our ability to improve: a call for investment in a new future*. 2003. Disponível em: <http://www.almaden.ibm.com/coevolution/pdf/engelbart_paper.pdf>. Acesso em 26 jun. 2013.

FEENBERG, A. Do essencialismo ao construtivismo. In: NEDER, R. T. (Org.). *A teoria crítica de Andrew Feenberg: racionalização democrática, poder e tecnologia*. Brasília: Observatório do Movimento pela Tecnologia Social na América Latina; CDS; UnB; Capes, 2010. p. 203-251.

FIGUEIREDO, N. M. de. *Estudos de uso e usuários da informação*. Brasília: IBICIT, 1994.

INGWERSEN, P. Conceptions of Information Science. In: VAKKARI, P. C. B. (Ed.) *Conceptions of library and information science: historical, empirical and theoretical perspectives*. London: Taylor Graham, 1992. p. 299-312.

MORIN, E. Da necessidade de um pensamento complexo. In: MARTINS, F. M.; SILVA, J. M. (Org.). *Para navegar no século XXI*. Porto Alegre: Sulina; Edipucrs, 2003. p.13-39.

RAYWARD, W. B. Visions of Xanadu: Paul Otlet (1868-1944) and Hypertext. *Journal of the American Society for Information Science*, v. 45, n. 4, p.235-250, 1994. Disponível em: <http://people.lis.illinois.edu/~wrayward/Visions%20of%20Xanadu_JASIS.pdf>. Acesso em: 15 set. 2014.

SARACEVIC, T. Ciência da Informação: origem, evolução e relações. *Perspectivas em Ciência da Informação*, Belo Horizonte, v. 1, n. 1, p. 41-62, Jan./Jun. 1996.

SMIT, J. W. et al. A determinação do campo científico da Ciência da Informação: uma abordagem terminológica. *Datagramazero - Revista de Ciência da Informação*, Rio de Janeiro, v. 5, n. 1, fev. 2004. Disponível em: <http://www.dgz.org.br/fev04/Art_03.htm>. Acesso em: 17 abr. 2013

TAN, P.; STEINBACH, M.; KUMAR, V. *Introdução ao DATAMINING mineração de dados*. Rio de Janeiro: Ciência Moderna Ltda., 2009.

WERSIG, G. Information science: the study of postmodern knowledge usage. *Information Processing and Management: an International Journal*, Tarrytown-Nova Iorque, v. 29, n. 2, p. 229-239, Mar./Apr. 1993.