

INFORMAÇÃO, FONTES, WIKIPÉDIA: QUESTÕES LEVANTADAS: APONTAMENTOS NECESSÁRIOS¹

Jackson da Silva Medeiros²

Universidade Federal do Rio Grande do Sul

jackson.medeiros@ufrgs.br

Rodrigo Silva Caxias de Sousa³

Universidade Federal do Rio Grande do Sul

rodrigo.caxias@ufrgs.br

Resumo

Problematiza questões relativas ao uso da Wikipédia como fonte de informação. Expõe diferentes noções de enciclopédia, considerando funções, características e alternativas de avaliação a partir de distintas contribuições teóricas. Apresenta a Wikipédia como um modelo de construção e disseminação do conhecimento, observando seus princípios e a diversidade de perspectivas quanto à confiabilidade das informações identificadas entre seus verbetes. Conclui que a Wikipédia se materializa através de um modelo próprio de construção e distribuição de conhecimento, pautada por uma lógica na qual a inteligência coletiva se sobressai à noção de credibilidade validada por especialistas, característica das enciclopédias tradicionais.

Palavras-chave: Enciclopédia. Fonte de informação. Wikipédia. Validade da informação.

INFORMATION, SOURCES, WIKIPEDIA: ISSUES RAISED: NECESSARY NOTES

Abstract

The use of Wikipedia as an information source. Exposes different notions of encyclopedia, considering functions, characteristics and alternatives of evaluation from different theoretical contributions. It presents Wikipedia as a model of construction and dissemination of knowledge, observing its principles and diversity of perspectives as to the reliability of the information identified between its entries. Concludes that Wikipedia materializes itself through an own model of construction and distribution of knowledge, guided by a logic in which collective intelligence excels at the notion of credibility validated by specialists, characteristic of traditional encyclopedias.

Keywords: Encyclopedia. Information source. Wikipedia. Validity of information.

¹ Trabalho com financiamento do CNPq (Processo - 431367/2016-7).

² Doutor em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul
Professor do Departamento de Ciências da Informação da mesma Universidade

³ Doutor em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul
Professor do Departamento de Ciências da Informação da mesma Universidade



1 INTRODUÇÃO

Há muito tempo é notável o acúmulo de conhecimento que circunda nossas vidas. Cada vez mais a informação se torna essencial para tomadas de decisões e para construção de novos conhecimentos. Nessa lógica, as fontes de informação são canais vitais para a busca e a consulta a conteúdos sobre determinado assunto, principalmente quando disponibilizadas na internet, proporcionando caráter de atualização constante. Isso ocorre porque a possibilidade de criação de conhecimento de forma coletiva, reunida em repositórios de informação, fomenta as formas de saber e conhecer da sociedade.

No entanto, essa forma de enxergar o conhecimento, baseada na colaboração e na diversidade é desafiadora, uma vez que infunde em um consulente, seja ele um profissional ou um leigo, processos que vão além da mera localização de conteúdos informacionais. A imensa disponibilidade de conteúdo online faz com que haja necessidade de capacidade crítica para identificar potenciais fontes confiáveis de informação para que suas necessidades sejam atendidas.

Já no final dos anos 1990, Campello, Caldeira e Macedo alertavam sobre as mudanças da informação sofre em questões de construção, disponibilização etc., influenciando nisso a quantidade e qualidade das formas de conhecimento, informação e saber que são disponibilizadas aos sujeitos, tanto os que desejam comunicar algo quanto os receptores desse algo. A Wikipédia faz parte desse tipo de iniciativa sobre as formas de conhecer e compartilhar informações sobre os mais diversos assuntos.

Este artigo foi pensado a partir de trabalho apresentado no I International Wikipedia Scientific Conference e no II Brazilian Wikipedia Scientific Conference e das discussões lá propostas. Durante a conferência, diversos debates foram levantados, em especial sobre a utilização da Wikipédia como fonte de informação. O que se pôde depreender é que ainda existe grande confusão conceitual sobre a utilização do conteúdo da plataforma, levando à discussão (que é objetivo deste trabalho) sobre sua validade e se o tipo de informação ali existente pode ou não ser considerada científica.

Como ponto de partida, este trabalho busca estabelecer um pano de fundo dessas questões, realizando uma exploração da Wikipédia a partir de abordagem de definições/caracterizações/noções/ideias sobre o que vem a ser uma enciclopédia e sua utilização como fonte de informação, principalmente com aportes das áreas de Biblioteconomia e Ciência da Informação. A partir de caráter ensaístico, apresenta metodologia qualitativa e abordagem bibliográfica não exaustiva para alicerçar e proceder a proposta.

2 INFORMAÇÃO E (CLASSIFICAÇÃO DE) FONTES

De maneira genérica, podemos entender que a informação depende do entendimento, pela recepção, da significação dos acontecimentos. É, por isso, uma constituição de um saber que se acumula e pode ser comunicado. Essa proposição pode nos ajudar a pensar as fontes de informação no sentido das informações que abarcam, uma vez que a diversidade dessas fontes e a manutenção de suas características implica em processos de acumulação de conteúdo, informação, conhecimento etc. que demonstram complexidade em todo seu fluxo, admitindo a dificuldade de definição do termo.

Por esse viés, a informação depende de um processo de recepção, de um “fazer sentido”, coordenado pelas atividades que estão além daquele que a produz, mas a partir das necessidades e das possibilidades exploratórias daquele que recebe e do modo como a manipula. Há diversas forças que coordenam o que pode ser definido por informação, gerando possibilidades de compreensão do conceito. Como reflete González de Gómez (2009), a informação só pode ser colocada, pensada, conceituada, entendida a partir da sua história social, onde o “ser” informação só pode ser compreendido com base no por que, para que e por quem é outorgado nas ordens do ser ou do saber.

Se há relação que o reconhecimento da informação depende de intencionalidade de um receptor e de um produtor, entra em cena um círculo que trabalha em torno da confiabilidade da informação. Por esse ponto é interessante notar como isso é percebido para a ciência, já que, como declara Mueller (2000, p. 22), neste campo, a confiabilidade é uma de suas características mais importantes. Para a autora, a correção de uma informação, isto é, o que credita informação científica como confiável está ligada ao “conhecimento obtido segundo uma metodologia científica, ou seja, é o resultado de pesquisas realizadas por cientistas, de acordo com regras definidas e controladas”. A autora descreve que esse processo, ao passar também pelos pareceres e pela divulgação, distingue-se do conhecimento não científico, comum, uma vez que mantém, pela legitimação, “o processo pelo qual o ‘legislador’ encarregado de zelar pelo discurso científico é autorizado, pela comunidade científica, a prescrever as condições que estabelecem se determinado conhecimento pode ser considerado científico.” (MUELLER, 2006, p. 30).

Para Carrizo Sainero (1994) é possível considerar um documento como fonte de informação quando é autêntico, confiável e acessível. No primeiro caso, comprovando sua origem, estabelecendo dados objetivos, como autoria, edição etc. O segundo diz respeito a

verificabilidade da informação, baseado em autores e/ou editores qualificados e com prestígio reconhecido. É acessível quando está disponível para consulta e uso. Dessa forma, embora o conceito de fonte de informação seja amplo, é possível considerá-lo como “materiais ou produtos, originais ou elaborados, que abarcam notícias ou testemunhos através dos quais se tem acesso ao conhecimento, qualquer que seja.” (CARRIZO SAINERO, 1994, p. 30).

Fontes de informação são classificadas, tradicionalmente, em primárias, secundárias e terciárias. Fontes primárias são aquelas publicações que contêm dados novos ou originais ou, ainda, novas interpretações de ideias, abarcando periódicos, publicações científicas ou oficiais, patentes, normas, teses, entre outras. As secundárias tomam as primárias como insumo, organizando-as a partir de esquemas próprios, ou seja, um tipo de interpretação ou condensação das primeiras, constituindo produtos elaborados não mais pelos publicadores das fontes primárias, mas a partir de um tratamento que essas recebem, como, por exemplo, índices, manuais, enciclopédias e fontes históricas. Para as fontes terciárias, o comum é admiti-las como um guia das fontes primárias ou secundárias, não estabelecendo conhecimento, mas apenas a indicação sobre documentos primários ou secundários, como o caso de uma bibliografia de bibliografias (CARRIZO SAINERO, 1994; CUNHA, 2001; GROGAN, 1973).

Outra classificação comumente feita em se tratando não mais do tipo de informação, mas dos canais de comunicação da informação, principalmente no que se refere à informação científica, são os nomeados canais formais ou informais. Os primeiros mantêm amplo acesso e estão vinculados aos estágios mais avançados de estudos, como periódicos e livros; os segundos canais dizem respeito à atividade considerada pessoalizada, compreendendo pesquisa inconclusas e algumas apresentações em congressos (MUELLER, 2000).

Grogan (1973) ainda busca categorizar as fontes não-documentárias, aquelas que caracteristicamente vão fluir por canais de comunicação específicos, como as comunicações institucionais, de sociedades profissionais, de indústrias, de universidades etc., e as informais, a partir de colegas, visitantes e até mesmo conversas de corredor. O próprio autor aponta a importância, não sem discorrer por possíveis problemas de confiabilidade, das fontes informais de informação, descritas como fontes “vivas”, visto que não se pode menosprezar a relevância dessas fontes não-documentárias de informação. Formas de conhecimento que surgem a partir desse tipo de fontes são as conferências e os colégios invisíveis. Enquanto as primeiras podem ser vistas apenas pelo lado das comunicações publicáveis, outro lado pode ser explorado a partir das redes de investigadores que se formam em torno de um assunto e de como se organizam e articulam suas relações em eventos que se intercalam com as comunicações/apresentações. Da mesma forma os colégios invisíveis são onde contatos são postos em prática e podem mostrar

uma troca de mensagens constante com finalidades diversas. Esses colégios são formados, em geral, por investigadores com titulação formal e atividade científica em uma área.

Embora a classificação tradicionalmente estabelecida seja teoricamente bem ajustada, na prática, ao pensar a associação das tipologias de informações que podem existir em uma fonte de informação, suas delimitações podem se tornar difusas (GROGAN, 1973). Buscar uma classificação das fontes de informação pode ser uma tarefa inútil, visto que as categorizações propostas são amplas e, por mais que busquem encerrar um determinado mundo documental, deixam – e não poderia ser diferente – que suas fronteiras sejam fluidas, afetando a própria classificação. Estamos convencidos, no entanto, que uma classificação, além de ser complexa e levar em consideração fatores, muitas vezes, subjetivos, tende a cair em certo desuso, uma vez que a informação disponível na web é criada e passa a compor elementos que, em muitos casos, diferem da informação estática disponível, por exemplo, em livros.

2.1 ENCICLOPÉDIA COMO FONTE DE INFORMAÇÃO

O que temos percebido, principalmente em relação à literatura de Biblioteconomia e Ciência da Informação, diz respeito ao tratamento da enciclopédia como fonte de informação ainda no formato clássico, geralmente pensado no modo impresso ou, quando digital, mantendo as tradicionalidades do formato. Entendemos que seja pensado dessa maneira pois, nessa configuração, características de autoridade e de confiabilidade podem ser mais claramente definidas.

Por outro lado, a expansão da Wikipédia na última década, somando esforços de especialistas e não especialistas, deve ser considerada em análises que tomam enciclopédias como fonte de informação, ainda que esse seja um processo incipiente. Nestas seções (2.1 e 2.2), nossa ideia é explorar alguns conceitos sobre a como a literatura tem tratado a enciclopédia e buscar situar a Wikipédia nesse contexto.

O Dicionário Houaiss diz que uma enciclopédia é, ambiciosamente, o “conjunto de todos os conhecimentos humanos” (HOUAISS, 2009), caracterizando-se por ser uma “obra que reúne todos os conhecimentos humanos ou apenas um domínio deles e os expõe de maneira ordenada, metódica, seguindo um critério de apresentação alfabético ou temático”. Grogan (1973) nos parece claro ao afirmar que uma enciclopédia não traz uma palavra final sobre algo no mundo, mas os fatos primeiros e essenciais, visando a facilidade de encontrar a informação que se procura, não primando pela atualidade ou pela exaustividade do assunto, uma vez que

contém informações específicas ou orientações sobre alguma temática. Caracteriza-se, assim, por oferecer um ponto de partida para investigações mais aprofundadas.

Para a Biblioteconomia e a Ciência da Informação, como apontamos, a enciclopédia em formato de livro ainda é tomada como um padrão, inclusive para avaliação. Já nos anos 40, Flexner (1943, p. 97) apontava a enciclopédia como um livro⁴ de referência que visa “responder a uma questão específica no menor tempo possível e da maneira mais direta [...] diferindo de livros [...] não necessariamente no assunto mas na quantidade de informação e na organização e apresentação do material.”. A posição de considerar enciclopédias como livros de referência também é abordada por Grogan (1973).

Para Cunha (2010), uma enciclopédia é, por excelência, uma fonte de informação onde o conteúdo de artigos⁵ são desenvolvidos por especialistas de forma curta e condensada, podendo incluir algum tipo de classificação e trazendo ilustrações, gráficos e tabelas, bem como referências sobre o assunto tratado. No entanto, não deve ser tomada como única fonte confiável para a execução de um trabalho ou, ainda, para conhecimento sobre determinado assunto. Flexner (1943) age na mesma toada ao afirmar que uma boa enciclopédia contém artigos que são escritos por especialistas em seus respectivos campos e não por indivíduos que podem cobrir genericamente qualquer área, ressaltando que apenas um especialista é capaz de conhecer a importância do assunto que está sendo tratado/apresentado.

No que tange as funções de uma enciclopédia, para Cunha (2010), são as que seguem: (i) fornecer respostas factuais ou imediatas (quem, como, qual, quando, onde?); (ii) dar informações sobre antecedentes ou retrospectivas; e (iii) sugerir leituras complementares. Uma enciclopédia, é, nesse sentido, a mais importante das obras de referência tendo em vista que: (a) pela amplitude de seu objetivo permite ao leitor fazer busca e encontrar respostas para as mais variadas indagações; (b) por ser informativa provê os dados básicos sobre assuntos diversos, muitas vezes redirecionando o leitor para outros tipos de fontes de informação; (c) por meio de seus anuários ou suplementos a enciclopédia se atualiza, aumentando seu ciclo de vida; e (d) pode atender a uma variedade de tipos de leitores.

Flexner (1943) aponta uma série de características das enciclopédias, como listadas: (a) contém informações sobre variados campos de conhecimento; (b) encaminhamento, através de informação e bibliografias, para coleções mais amplas de documentos; (c) indicação de material

⁴ Embora Flexner faça sua discussão a partir do livro, e não do documento, entendemos que isso reflete um período que descreve a importância dada a esse artefato e seu resgate pode ser importante para o nosso trabalho.

⁵ Neste trabalho, utilizamos os termos artigos e verbetes como sinônimos para identificação do conteúdo de enciclopédias.

de referência ao final dos artigos; (d) organização para uso rápido e fácil; (e) constituem importante fonte para os iniciantes em um assunto, provendo simples introdução, boas definições, ponto(s) inicial(is) para investigação mais profunda e rápida revisão a pontos específicos já familiares.

A autora também apresenta, de forma pragmática, alguns elementos para avaliação de uma enciclopédia, como por exemplo, como saber se uma enciclopédia responde questões de assuntos específicos, indicando que se deve realizar testes no campo que o consulente melhor conhece. Também fala da simplicidade de uso, aconselhando realizar a consulta em campos ou assuntos pouco conhecidos, visando verificar se ela fornece uma boa introdução ao assunto ou se encaminha para elementos mais amplos e/ou complexos. Isso permite considerar que, segunda a autora, “Testar o valor de um livro de referência envolve considerar sua utilidade na relação com o propósito pretendido.” (FLEXNER, 1943, p. 97). Com isso, cumpre elaborar que, enquanto designadas como tais, enciclopédias se caracterizam como algo próximo a uma sistematização de conhecimentos ordenados de forma concisa para consulta rápida e fácil, proporcionando vários níveis de profundidades nas abordagens dos assuntos que trata.

Esse formato mais tradicional sofreu grande choque a partir da popularização de sistemas computacionais e da forma como as pessoas lidam com informação em ambiente digital. Lidar com informação, neste caso, passa a ser uma ampla designação, já que começa a atuar em diversos aspectos do fluxo informacional, fazendo com que o usuário leitor, receptor de informação, consulente de enciclopédias tradicionais ganhe a possibilidade de se manifestar e construir algo que pode, em certa medida, servir de subsídios para outros indivíduos.

Provavelmente o início desse processo tenha sido visto nos blogs dos anos 1990 – naquela época ainda como espécie de diários online – e se redefinido a partir do momento em que se enxergou essa plataforma como um modo de criar conteúdo, de disseminar informações, de sublinhar pontos de vista sobre assuntos variados, registrar um viés de quem escreve e uma percepção daquele que responde (em comentários, por exemplo). Esse processo foi criando uma sistemática de colaboração entre os indivíduos, com respostas rápidas e construções que não mais eram estabelecidas à poucas mãos.

Esse tipo de conteúdo não mais criado por indivíduos que tinham por base uma autoridade ou expertise no seu campo fez com que autores, como Tomaél e colaboradores (2004), buscassem apontar e/ou estabelecer critérios de qualidade para avaliar fontes de informação na internet, baseando-se, principalmente, na qualidade da apresentação e do conteúdo dessas fontes. No que diz respeito à forma, os critérios estão ligados às questões de arquitetura, usabilidade e disponibilidade da informação. No que tange ao conteúdo, muitos

critérios de avaliação estão vinculados ao ato de verificar credenciais ou autoridades pessoais e institucionais, permitindo que isso defina, a partir de corpos editoriais, a credibilidade da informação que a fonte disponibiliza. Esses critérios estão baseados, conforme os autores, na tentativa de garantir um contraponto à alta volatilidade das informações em ambiente web.

Após análise e teste piloto, os autores buscam definir critérios de qualidade na avaliação, como na identificação, consistência e confiabilidade das informações, adequação da fonte, links, facilidade de uso, layout, restrições percebidas e suporte ao usuário. Anotando que esse trabalho data de 2004, ainda não tinha recebido o impacto da Wikipédia como fonte de informação utilizada em larga escala, atrelando suas análises, ainda, aos materiais disponíveis em formatos tradicionais ou opinativos, como os próprios blogs de caráter pessoal ou jornalísticos, por exemplo.

A diversidade de fontes de informação existentes na internet praticamente inviabiliza que sejam elencados elementos estáticos de avaliação, o que torna o processo de medição de confiabilidade uma questão que envolve determinados parâmetros suscetíveis à análise não apenas da informação, mas também da ferramenta. Isso é corroborado por Braz e Souza (2014, p. 22) ao afirmarem que “os critérios não podem ser generalizados para todos os endereços”.

77

A relação do saber com as enciclopédias, podemos dizer, mudou radicalmente em especial a partir da criação da Wikipédia. Essa alteração se deu em vários aspectos, como a disponibilidade, acesso, uso, mas, de forma primordial, na maneira como o conhecimento ali depositado é construído e constituído. Buscamos aclarar isso na próxima seção.

2.2 WIKIPÉDIA COMO UM TIPO DE ENCICLOPÉDIA: MODELO, CRÍTICAS E FORMAS DE EXISTIR⁶

Nascida em 2001, a Wikipédia é uma ferramenta, em forma de enciclopédia, que possibilita a disseminação de conteúdos, sendo a criação e a edição dos artigos nela contidos realizadas por diversos indivíduos com enorme variedade de conhecimento sobre o assunto. Sua construção ocorre de forma colaborativa e seu conteúdo multilíngue está baseado em licença livre⁷, possibilitando que o conhecimento das multidões seja elaborado e depositado em ambiente propício para tal. Segundo o site Alexa, fornecedor de dados sobre tráfego em websites de todo o mundo, a Wikipédia está na quinta posição dessa lista⁸, responsável por,

⁶ Não será realizado um resgate histórico da Wikipédia e como se dá o processo de construção de verbetes. Nossa intenção é pensar sobre as formas de existência da Wikipédia enquanto fonte de informação.

⁷ Todos os artigos da Wikipédia estão sob licença GNU Free Documentation License.

⁸ <https://www.alexa.com/topsites> em 18 de junho de 2018.

aproximadamente, 65% do tráfego gerado a partir de buscadores. Como fonte de informação, hoje, nada é comparável à Wikipédia. A enciclopédia está disponível em centenas de línguas, alcançando, em 2017, mais de 980 mil artigos em português e mais de 5,5 milhões de artigos em inglês.

O modelo em que a Wikipédia se baseia para a publicação de conteúdos está ligado à construção coletiva que a plataforma proporciona, sendo afetada por eventos cotidianos que adquirem caráter de referência, como, por exemplo, notícias em sites de jornalísticos ou sobre celebridades. Isso mostra dois pontos em polos distintos: o primeiro é a celeridade com que os conteúdos são atualizados, possibilitando que as informações ali contidas mantenham traços de correção praticamente em tempo real. Este afã de atualização constante marca o outro polo, uma vez que a corrida por atualizações fundamentadas pode trazer problemas para a credibilidade dos artigos.

Com relação a isso, a Wikipédia se encaixa na classificação de fonte terciária de informação, uma vez que se coloca como uma espécie de guia sobre conhecimento estabelecido, não atuando como geradora de conhecimento⁹. Toma por base elementos já publicados em “canais confiáveis”, como artigos de periódicos com revisão por pares, sites de grandes corporações jornalísticas ou de entretenimento e atividades de páginas pessoais de indivíduos ou organizações reconhecidas com autoridade em um assunto ou campo. A Wikipédia mantém uma página que trata sobre a identificação de fontes confiáveis¹⁰ e também descreve como a plataforma trabalha com a verificabilidade¹¹ delas¹².

A Wikipédia baliza seus princípios em cinco pilares¹³, a saber: enciclopedismo, neutralidade de ponto de vista, licença livre, boa convivência e liberalidade nas regras. O primeiro está ligado, obviamente, à sua autodeclaração como uma enciclopédia, baseando-se tanto nas características quanto nas políticas que a definem como tal. O segundo item está fundamentado na imparcialidade de pontos de vista, isto é, a existência de visões sobre alguns assuntos deve ser refletida no artigo, evitando julgamentos sobre pontos melhores ou piores, o que implica na referência às fontes utilizadas. O terceiro item, referente ao licenciamento do

⁹ https://en.wikipedia.org/wiki/Wikipedia:No_original_research

¹⁰ https://en.wikipedia.org/wiki/Wikipedia:Identifying_reliable_sources

¹¹ <https://en.wikipedia.org/wiki/Wikipedia:Verifiability>

¹² As referências da Wikipédia utilizadas neste trabalho são da sua versão inglesa. Essa decisão foi tomada devido a este ser o idioma mais falado no mundo e, por consequência, as contribuições em inglês que a enciclopédia recebe podem ser consideradas mais consistentes devido, também, ao número de possível editores que funcionam como filtro das informações ali contidas. Em caráter complementar, o próprio artigo sobre “confiabilidade” declara que seu conteúdo parte de uma análise, predominantemente, da versão inglesa da enciclopédia.

¹³ https://en.wikipedia.org/wiki/Wikipedia:Five_pillars

conteúdo, possibilita que qualquer indivíduo edite artigos sobre qualquer assunto, não atribuindo propriedade sobre as contribuições. O quarto elemento diz respeito à boa conduta dos atores envolvidos na edição da enciclopédia e os chamados “comportamentos tóxicos”, como desrespeito por qualquer circunstância. O último ponto relata a não existência de regras fixas, uma vez que isso reflete a própria ideia da Wikipédia, ou seja, uma construção conjunta e formação de conhecimento, inclusive das regras, em tempo real.

Na Wikipédia não existem autores individuais. Aliás, podemos pensar que a terminologia “autor” nem mesmo consta na ideia da Wikipédia, uma vez que a autoria é coletiva, colaborativa, onde ninguém é o dono, o detentor do conhecimento. Pensar o coletivo, e em especial suas ações colaborativas, age na organicidade do conceito e, por conseguinte, da própria Wikipédia como instrumento de construção colaborativa de conhecimento a partir das suas características como enciclopédia livre baseada em contribuições voluntárias. Esse formato pode ser pensado a partir da ideia de sabedoria das multidões.

Na obra *A sabedoria das multidões*, Surowiecki (2006) aponta alguns elementos teóricos e pragmáticos para constituir sua conceituação. O autor elenca quatro elementos fundamentais para que isso ocorra: (a) diversidade de opiniões, onde cada agente oferece/disponibiliza informações pessoais; (b) independência, onde é possível agregar informações/opiniões sem constrangimentos dependentes de interesses; (c) descentralização, possibilitando trabalho e especialização com conhecimento local; e (d) agregação com um tipo de mecanismo que fomente a organização desorganizada de ideias para transformar esse processo em uma decisão coletiva. Essas questões, defende Surowiecki (2006, p. 26), apresentam-nos que “A forma mais simples de conseguir boas respostas confiáveis é perguntar ao grupo a cada vez”, uma vez que um grupo independente, ao fornecer algum tipo de informação, faz com que os erros sejam anulados entre si.

A diversidade é um ponto importante a ser considerado na colaboração, uma vez que ela acrescenta possibilidades que de outra forma estariam ausentes, permitindo que aspectos irrelevantes ou inconsistentes das decisões grupais sejam cancelados. Ou seja, o desempenho do grupo depende menos dos especialistas inteligentes e mais da diversidade no nível do grupo; inteligência sozinha não garante diferentes perspectivas proporcionadas pela diversidade. Para a consolidação dessa sabedoria, os caminhos a serem trilhados são o envio do maior número de indivíduos e opiniões possíveis à plataforma, fazendo com que o sistema possa encorajar e especular sobre o sucesso ou fracasso das informações fornecidas, seja na inserção ou na validação, garantindo ideias inovadoras e diferenças significativas entre elas. É um processo

claramente descentralizado, onde não há um núcleo coordenador das atividades a serem executadas (SUROWIECKI, 2006).

Existem, de outro lado, visões negativas sobre o assunto. Como dissemos, há em muitos casos, principalmente na Biblioteconomia e na Ciência da Informação, uma valorização de ferramentas tradicionais em detrimento de inovações. Cunha (2010, p. 4) estabelece uma crítica contra os instrumentos abertos proporcionados pela internet, como o Google e a Wikipédia, indicando uma vanguarda dos títulos impressos ou validados por especialistas, relacionando isso com a correção dos instrumentos. Segundo o autor, “A Wikipédia, construída por milhares de voluntários anônimos, apresenta, em grande parte devido a esse anonimato, a possibilidade de erros ou omissões”. Outra discussão, relacionada à Wikipédia, parece aberta na literatura da área de Ciência da Informação é colocada por Campello (2008). A autora não mostra dados relativos à sua afirmação, mas declara que a Wikipédia, baseada em um modelo de colaboração aberta e voluntária parece não sustentar. Nesse sentido não temos como elencar elementos que a autora considera pertinentes para sua consideração.

Assim também faz Andrew Keen no livro **O culto do amador: como blogs, MySpace, YouTube e a pirataria digital estão destruindo nossa economia, cultura e valores**. A obra, publicada na versão em inglês em 2007, mostra um autor preocupado com a projeção da existência de milhões de blogs em 2010. Estes, na visão de Keen, iriam corromper e confundir coletivamente a opinião popular, desvirtuando a linha entre o que é verdadeiro e o que é falso, levando a que os defensores desse modo de ver a informação agiriam como se toda postagem fosse a criação de uma verdade. Mas isso, como afirmar Keen, é uma preocupação antiga, visto que as maneiras de como se divulga ou se adquire informações são preocupações que apenas se potencializam com a utilização da internet.

No que tange à Wikipédia, Keen (2009, p. 9), afirma, ao analisar casos problemáticos enfrentados pela plataforma, a existência de uma “celebração do amadorismo”. Para chegar a sua assertiva, aponta a quantidade de artigos generalistas da plataforma que não foram avaliados: “nenhum deles”, afirma o autor, foi “editado ou atentamente examinado quanto à sua exatidão.”. Um dos fatores que o autor, aponta para esse “declínio da confiabilidade da informação que recebemos, distorcendo assim, se não corrompendo por completo, nosso debate cívico” (KEEN, 2009, p. 30) é a falta de separação entre o leitor e o escritor, entre o amador e o especialista. Keen (2009), inclusive, critica o criador da Wikipédia, Jimmy Wales, lançando a pergunta: se ele largou a faculdade, como sabe quem está certo?, afinal, na visão de Keen, todos precisam dos especialistas.

As formas de obtenção de informação atualmente sofrem algum desvio quando se passa a considerar a grande incidência de *fake news* veiculadas, principalmente, na mídia¹⁴. Na Wikipédia isso não é diferente, fazendo com que a enciclopédia poderá ter seu conteúdo afetado.

Isso tem levado a estudos, entre outros, que discutem o comprometimento moral de jornalistas na divulgação de notícias (BORDEN; TEW, 2007), os parâmetros entre o que é informação/notícia falsa ou real (MCBETH; CLEMONS, 2011) e, inclusive, métodos de investigação e detecção de *fake news* (CONROY; RUBIN; CHEN, 2015). Sabendo do grande impacto da mídia na vida das pessoas e que esse processo toma mais corpo a partir da internet e das redes sociais, falar de fontes de informação torna-se, por si só, um processo complexo.

Aliado a isso, é possível pensar a gravidade disso para a plataforma no momento em que diversos artigos da Wikipédia tomam como fonte de informação veículos de notícias de todo mundo. Ou seja, os próprios editores de conteúdo da Wikipédia assumem que determinadas fontes jornalísticas são confiáveis, usando-as para mostrar um ponto de vista ou algo que a mídia noticiou.

Com constantes indagações e debates sobre a qualidade do conteúdo da Wikipédia, diversos estudos já foram conduzidos e a compararam em qualidade à outras enciclopédias, como a Britânica (GILES, 2005). Alguns autores (HU et al., 2007) ainda listam questões que dificultam o trabalho com a Wikipédia, como o grande número de artigos com diversas coberturas sobre assuntos de atualização constante e a variação de *background* dos colaboradores, trazendo informações subjetivas, além de problemas com *hoaxes*¹⁵.

Apesar de todos exageros alegados por todos os lados, tanto daqueles que ferrenhamente defendem que a inteligência coletiva da sabedoria das multidões, formada pelo conhecimento de não especialistas, pode se constituir como um campo iluminado a oferecer todas as respostas baseadas na quantidade de indivíduos e seus conhecimentos e diversidades, quanto dos defensores de que apenas especialistas devem se manifestar sobre informações de um campo, alumbrando mentes indoutas, passam ao largo de nossa análise. Como dissemos, nossa ideia é problematizar a informação disponível na Wikipédia a partir de seu conteúdo. Sabemos que essa tarefa é difícil e, de fato, constitui-se em algum tipo de incompletude. Buscamos, a seguir, elencar alguns desses fatores.

¹⁴ Chamamos “mídias”, de maneira genérica, as tradicionais formas de acesso às notícias, como jornais e revistas com certa notoriedade, mas que hoje atuam também de forma bastante vasta na internet.

¹⁵ Boato ou trote, em tradução livre.

3 WIKIPÉDIA COMO FONTE DE INFORMAÇÃO: QUESTÕES QUE SE LEVANTAM, APONTAMENTOS QUE SE FAZEM NECESSÁRIOS

Pelo seu caráter de colaboração aberta, como dissemos, a Wikipédia assume que qualquer indivíduo pode editar seu conteúdo. Isso traz, sem dúvida, algumas questões que transitam ao redor da sua confiabilidade. A própria enciclopédia possui um artigo chamado *Reliability of Wikipedia*¹⁶ (Confiabilidade da Wikipédia), informando que isso tem sido testado por meios estatísticos através de revisão comparativa, análise dos padrões históricos e pontos fortes e fracos referentes ao processo de edição da enciclopédia e que, em muitos casos, chama atenção da mídia e de investigadores.

A Wikipédia também referencia no verbete uma série de estudos que foram conduzidos com a intenção de comparar a confiabilidade da enciclopédia online com outras enciclopédias conhecidas tradicionalmente, como a Britânica. A análise de verbetes específicos, como patologia, psiquiatria, oncologia, entre outros, mostrou que a Wikipédia consegue uma grande cobertura dos assuntos, gerando, inclusive, alta taxa de confiabilidade no conteúdo disponibilizado. No mesmo artigo, por outro lado, a enciclopédia relata casos de conteúdos que foram adicionados a artigos com intuito difamatório, piadístico ou ainda a criação de histórias ou personagens que nem mesmo existiam. Em alguns casos, inclusive, esses artigos com informações, diríamos, controversas, foram replicadas até mesmo em livros universitários.

No entanto, é importante observar que o trabalho da coletividade traz, através da diversidade, a agregação de indivíduos mais ou menos *experts* em determinado conteúdo, o que pode se mostrar fundamental na exploração de alternativas ou pontos de vista (SUROWIECKI, 2006), demonstrando que isso é fundamental para a constituição da Wikipédia. Sua abertura pode, até mesmo, fazer com que essa sabedoria seja utilizada para envolvimento de voluntários e desenvolvimentos de softwares de correção de conteúdo.

Mas tomemos como exemplo os próprios pontos levantados por Surowiecki. Em suas argumentações, o autor se utiliza de estudos conduzidos por especialistas em diversos campos para justificar ou explicar suas propostas, ou seja, utiliza-se de informações cientificamente validadas para elaboração de reflexões em seus estudos. Isso é um retrato da condição da informação, principalmente científica, a qual permite aos experts lançarem mão de uma série de conhecimentos que se aplicam, muitas vezes, a um (micro)campo dentro da área, fazendo que decisões sejam baseadas no conhecimento que o próprio campo constrói e distribuiu aos seus membros.

¹⁶ https://en.wikipedia.org/wiki/Reliability_of_Wikipedia

Há nessa questão algo que o próprio Surowiecki aponta, mas não desenvolve (até mesmo porque fugiria ao objetivo de sua obra): o aparelho de informação científica, com toda sua estruturação hierárquica de pessoal e de ideias, pode ser considerado como baseado em um tipo de herança. Isso alicerça as bases de conhecimento, sejam elas cognitivas ou relacionais entre os indivíduos. Ou seja, os herdeiros de uma cultura de conhecimento acabam por fechar o círculo de conhecimento existente sobre um assunto. Em outras palavras, podemos mencionar certa incongruência na defesa do ponto de vista, mas, sabendo da necessidade de comprovação por “autoridades”, é algo que se faz necessário.

Podemos chegar a alguns elementos nessa relação: a Wikipédia desenvolveu, ao longo dos últimos anos, sua própria credibilidade na web. Em geral, usuários não especializados parecem tender a confiar, sem críticas, ao conteúdo disponibilizado nessa fonte de informação. Isso acontece também, muitas vezes, com veículos de mídia que não possuem profissionais especialistas em determinado assunto, pautando suas ações, algumas vezes, pelos assuntos “quentes” do momento na Wikipédia. Existem, inclusive, ferramentas que disponibilizam esse tipo de material, como o Wikipedia Live Monitor¹⁷, que acompanha os assuntos com alta taxa de edição e indica possíveis *breaking news*.

Isso leva a algumas questões que consideramos anteriores a toda problemática sobre a confiabilidade do conteúdo da Wikipédia: a falta de verificação das informações apresentadas, acarretando ou em informação construída de maneira errônea ou em reprodução desse tipo de informação; algo potencializado pelas redes sociais e pela velocidade com que a mídia trabalha na divulgação de material, principalmente, online. Consonante a isso, algumas interrogações se levantam: Como verificar as informações disponíveis na Wikipédia? Há real interesse em investigar isso? Como os (auto)denominados profissionais da informação lidam com essas questões? Quais métodos são utilizados para validar informações disponíveis na Wikipédia? Como orientar usuários a verificar a confiabilidade do conteúdo da Wikipédia? As questões são muitas e, pelo menos neste momento, não é nosso intuito desvelá-las, mas, sim, trazer a lume a complexidade que envolve o tratamento da informação em ambiente colaborativo na web.

Recompondo essas indagações, talvez o que deva ser considerado está além do processo colaborativo em si, uma vez que este é um modelo adotado pela Wikipédia e um dos elementos que a permitiu se constituir na maior enciclopédia online existente. Poderíamos nos perguntar, neste sentido, qual a taxa, ou ainda, com que rapidez ocorrem as correções de conteúdos inverídicos ou com algum tipo de problema na Wikipédia? Ou ainda, como são aprimorados os

¹⁷ <http://wikipedia-live-monitor.herokuapp.com> - Neste site é possível encontrar o *paper* escrito pelos pesquisadores que mostra os aspectos da pesquisa que constituem a ferramenta.

métodos de correção, sejam eles manuais, através dos colaboradores espalhados por todos o mundo, ou automáticos, através de softwares que automatizam o processo, deixando de lado a crítica ao modelo já adotado?

Nisso a Wikipédia atua, normalmente, de duas maneiras. A primeira através dos próprios editores, sejam eles usuários colaboradores ou com permissões especiais, para reverter alterações feitas com caráter de vandalismo. A outra maneira é desenvolvendo softwares que verificam os dados existentes na enciclopédia, podendo ir desde a aferição da qualidade de artigos da Wikipédia por modelos quantitativos e qualitativos que possibilitem identificação de bons artigos, verbetes com problemas identificáveis por aprendizagem de máquina, possibilitando contribuições naqueles que mantêm baixa qualidade e proporcionando melhora nos sistemas de busca (BLUMENSTOCK, 2008; CROSS, 2006; HU et al., 2007; WÖHNER; PETERS, 2009), até a identificação e reversão automática de problemas identificados, utilizando-se, por exemplo, avançados sistemas de aprendizagem de máquina (*machine learning*) (ADLER et al., 2011; CHIN et al., 2010; JAVANMARDI; MCDONALD; LOPES, 2011; POTTHAST; STEIN; GERLING, 2008; WEST; KANNAN; LEE, 2010).

Com isso nos juntamos a Cross (2006) ao apontar que o grande erro do entendimento da Wikipédia por parte daqueles que tecem críticas a seu respeito está em analisa-la com base nas expectativas de enciclopédias tradicionais. Adicionamos a isso, a ideia errônea de considerar o seu conteúdo como informação científica, uma vez que, embora sejam utilizadas fontes acadêmicas para construção de seus artigos, isso se constitui apenas em um fator de consulta às fontes, em tese, já verificadas. Ora, o propósito do conteúdo existente em uma enciclopédia já estabelece isso como algo infundado, já que, como vimos, enciclopédias pretendem fornecer uma visão geral e inicial de assuntos, e não sistematizações metodologicamente fundamentadas.

A base dos cinco pilares da Wikipédia nos parece dar um norte dessa abordagem, compreendendo que uma enciclopédia busca não uma neutralidade, mas a possibilidade de contemplar perspectivas discursivas em torno dos assuntos, através de informação construída por sujeitos que tomam por base fontes confiáveis já verificadas. Isso se alicerça em um senso de comunidade¹⁸ para desenvolvimento de conhecimento livre.

¹⁸ Estudos, como a tese de doutorado de Bernardo Esteves, intitulada *As controvérsias da ciência na Wikipédia em português: o caso do aquecimento global*, relatam as disputas, os conflitos dentro da Wikipédia. Como essa discussão não é objeto deste trabalho, indicamos a consulta à tese.

4 CONSIDERAÇÕES FINAIS

Os títulos dos livros de Denis Grogan (1973), *‘Science and Technology: an introduction to the literature’*, e de Murilo Bastos Cunha (2001), *‘Para saber mais: fontes de informação em ciência tecnologia’*, já dão uma pista de como a Biblioteconomia e a Ciência da Informação tratam o que consideram ser efetivamente fontes de informação. Isto é, a informação a ser tratada é aquela que é certificada, resultado de uma série de institucionalizações ligadas à ciência e que são capazes de substanciar aferições de credibilidade que o método científico fornece.

A Wikipédia é um modelo de uma enciclopédia, mas o ultrapassa, uma vez que agrega outros elementos em sua organização, constituição e distribuição de conteúdo. A figura do autor como fiador de credibilidade da informação é afastada em prol da inteligência coletiva que a sabedoria das multidões é capaz de fornecer.

Ao assumirmos a necessidade de uma autoridade formalmente instituída para a construção do conteúdo de uma enciclopédia, o foco do conteúdo é deslocado para a necessidade de instâncias autorizadas a dizer algo, a proferir um discurso, o que mais se caracteriza por restrições à utilização de determinadas fontes de informação; isso estaria ligado a um discurso científico com suas regras e formas de validade. Isso faria com que a enciclopédia fosse esquecida como objeto introdutório de informações, buscando essencialmente elementos de legitimação ou autoridade, nublando a condição da informação contida na Wikipédia.

Deve haver uma separação, mesmo que um modelo esteja balizado no outro, entre uma enciclopédia tradicional e o modelo enciclopédico utilizado pela Wikipédia, deixando de lado as críticas ao paradigma já instituído e amplamente utilizado na internet. Talvez sua forma de investigação se concentre na melhoria da qualidade dos artigos – algo que sempre pode ser refinado em qualquer enciclopédia. O interesse deve recair sobre as funcionalidades e os métodos de controle desse modelo já estabelecido e aceito, pelo menos fora do âmbito acadêmico.

Não se descarta, dessa forma, a utilização de conteúdos da Wikipédia como fonte de informação para construção de trabalhos científicos. O que não se deve amalgamar é a ideia de que a informação existente nos artigos da enciclopédia seja considerada informação científica, o que, definitivamente, não é seu papel. Ao compreendermos a Wikipédia como um tipo de enciclopédia – como ela se auto-intitulada –, é possível ter uma visão mais acurada de suas características e suas finalidades, colocando-a em um patamar de fonte de informação com um modelo próprio de construção e distribuição de conhecimento.

REFERENCIAS

- ADLER, B. T. et al. Wikipedia Vandalism Detection: Combining Natural Language, Metadata, and Reputation Features. INTERNATIONAL CONFERENCE ON INTELLIGENT TEXT PROCESSING AND COMPUTATIONAL LINGUISTICS, 2011, Tokyo, **Proceedings... CICLing 2011: Computational Linguistics and Intelligent Text Processing**, 2011. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-642-19437-5_23>.
- BLUMENSTOCK, J. E. Size matters: word count as a measure of quality on Wikipedia. INTERNATIONAL WORLD WIDE WEB CONFERENCE, 17., 2008, Beijing, **Proceedings...** . Disponível em: <http://www.jblumenstock.com/files/papers/jblumenstock_www08.pdf>.
- BORDEN, S. L. TEW, C. The Role of journalist and the performance of journalism: ethical lessons from “fake” news (seriously). **Journal of Mass Media Ethics**, v. 22, n. 4, p. 300–314, 2007. Disponível em: <<http://www.tandfonline.com/doi/abs/10.1080/08900520701583586>>.
- BRAZ, S. C. F.; SOUZA, E. D. Dos desafios da confiabilidade da informação na produção colaborativa de conteúdos: análises na Wikipédia, a Enciclopédia Livre. **Ciência da Informação em Revista**, v. 1, n. 3, p. 19-31, set./dez. 2014. Disponível em: <<http://www.seer.ufal.br/index.php/cir/article/view/1597/1191>>.
- CAMPELLO, B. S. Enciclopédias. In: _____; CALDEIRA, P. T. **Introdução às fontes de informação**. 2. ed. Belo Horizonte: Autêntica, 2008.
- CARRIZO SAINERO, G. Las fuentes de la información. In: _____; IRURETA-GOYENA SÁNCHEZ, P.; QUINTANA SÁENZ, E. L. **Manual de fuentes de información**. Madrid: CEGAL, 1994.
- CHIN, S. C. et al. Detecting Wikipedia vandalism with active learning and statistical language models. WORKSHOP ON INFORMATION CREDIBILITY, 4., 2010, Raleigh, North Carolina, USA. **Proceedings...** 2010. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.164.5660&rep=rep1&type=pdf>>.
- CONROY, N. J.; RUBUN, V. L.; CHEN, Y. Automatic Deception Detection: Methods for Finding Fake. **ASIST**, 2015, **Proceedings...** 2015. Disponível em: <<http://onlinelibrary.wiley.com/doi/10.1002/pra2.2015.145052010082/full>>.
- CROSS, T. Puppy smoothies: improving the reliability of open, collaborative wikis. **First Monday**, v. 11, n. 9, set. 2006. Disponível em: <<http://www.firstmonday.org/ojs/index.php/fm/article/view/1400/1318>>.
- CUNHA, M. B. **Para saber mais: fontes de informação em ciência e tecnologia**. Brasília: Briquet de Lemos, 2001.
- CUNHA, M. B. **Manual de fontes de informação**. Brasília: Briquet de Lemos, 2010.
- FLEXNER, J. M. **Making books work: a guide to the use of libraries**. New York: Simon and Schuster, 1943.

GILES, J. Internet encyclopaedias go head to head: Jimmy Wales' Wikipedia comes close to Britannica in terms of the accuracy of its science entries, a Nature investigation finds. **Nature**, v. 438, n. 15, dez. 2005. Disponível em: <<https://www.nature.com/articles/438900a.pdf>>.

GONZÁLEZ DE GÓMEZ, M. N. A reinvenção contemporânea da informação: entre o material e o imaterial. **Pesquisa Brasileira em Ciência da Informação**, v. 2, n. 1, p. 115-134, jan./dez. 2009. Disponível em: <<http://ridi.ibict.br/bitstream/123456789/116/1/GonzalezGomezTendencias2009.pdf>>.

GROGAN, D. **Science and Technology**: an introduction to the literature. 2. ed. rev. London: Clive Bingley: 1973.

HOUAISS, A. (Ed.) **Dicionário Houaiss Eletrônico**. 2009.

HU, M. et al. Measuring Article Quality in Wikipedia: Models and Evaluation. ACM CONFERENCE ON INFORMATION AND KNOWLEDGE MANAGEMENT, 2007, Lisboa, Portugal. **Proceedings...** Disponível em: <http://ink.library.smu.edu.sg/cgi/viewcontent.cgi?article=2515&context=sis_research>.

JAVANMARDI, S.; MCDONALD, D. W.; LOPES, C. V. Vandalism detection in Wikipedia: A high-performing, feature-rich model and its reduction through Lasso. INTERNATIONAL SYMPOSIUM ON WIKIS AND OPEN COLLABORATION, 7., Mountain View, CA, USA, 2011. **Proceedings...** Disponível em: <<http://www.pensivepuffin.com/dwmcphd/papers/Javanmardi.et.al-Vandalism-WikiSym11.pdf>>.

87

KEEN, A. **O culto do amador**: como blogs, MySpace, YouTube e a pirataria digital estão destruindo nossa economia, cultura e valores. Rio de Janeiro: Zahar, 2009.

MCBETH, M. K.; CLEMONS, R. S. Is fake news the real news? The significance of Stewart and Colbert for democratic discourse, politics, and polity In: AMARASINGAM, A. (Ed.). **The Stewart/Colbert Effect: Essays on the Real Impacts of Fake News**. Jefferson: McFarland, 2011.

MUELLER, S. P. M. A ciência, o sistema de comunicação científica e a literatura científica. In: CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. **Fontes de informação para pesquisadores e profissionais**. Belo Horizonte: UFMG, 2000.

_____. A comunicação científica e o movimento de acesso livre ao conhecimento. **Ciência da Informação**, v. 35, n. 2, p. 27-38, maio/ago. 2006. Disponível em: <<http://www.scielo.br/pdf/ci/v35n2/a04v35n2.pdf>>.

POTTHAST, M.; STEIN, B.; GERLING, R. Automatic Vandalism Detection in Wikipedia. EUROPEAN CONFERENCE ON INFORMATION RETRIEVAL, Glasgow, UK, 2008. **Proceedings...** 2008. Disponível em: <https://link.springer.com/chapter/10.1007/978-3-540-78646-7_75>.

SUROWIECKI, J. **A sabedoria das multidões**. Rio de Janeiro: Record, 2006.

TOMAÉL, M. I. et al. Critérios de qualidade para avaliar fontes de informação na internet. In: TOMAÉL, M. I.; VALENTIM, M. L. P. (Org.). **Avaliação de fontes de informação na internet**. Londrina: EDUEL, 2004.

WEST, A. G.; KANNAN, S.; LEE, I. Detecting Wikipedia vandalism via spatio-temporal analysis of revision metadata?. EUROPEAN WORKSHOP ON SYSTEM SECURITY, 3., Paris, 2010. **Proceedings...** 2010. Disponível em: <http://repository.upenn.edu/cgi/viewcontent.cgi?article=1461&context=cis_papers>.

WÖHNER, T.; PETERS, R. Assessing the quality of Wikipedia articles with lifecycle based metrics. INTERNATIONAL SYMPOSIUM ON WIKIS AND OPEN COLLABORATION, 5., Orlando, Florida, 2009. **Proceedings...** 2009. Disponível em: <<http://www.opensym.org/ws2009/Proceedings/p116-woehner.pdf>>.