



Qualidade na Wikipedia: revisão bibliográfica dos estudos sobre a enciclopédia livre¹

Franciane Tavares de MORAES²
Universidade Federal de Juiz de Fora, MG

Resumo

Desde que foi fundada, em 2001, a Wikipedia se tornou um fecundo objeto de estudo das mais diversas áreas da ciência. A análise da qualidade dos verbetes publicados é uma problemática frequente nas Ciências da Comunicação. Em vista disto, o presente artigo visa sistematizar e resumir os conceitos de diversos trabalhos, fornecendo uma revisão bibliográfica para a área e uma ferramenta inicial para uma pesquisa mais profunda sobre o tema.

Palavras-chave: Wikipedia; qualidade; Comunicação; revisão bibliográfica.

Introdução

A Wikipedia é uma referência exclusivamente online, escrita e editada por usuários, que segue o modelo de informação enciclopédico. Uma “enciclopédia livre”, cuja liberdade faz parte de todos os seus processos: o conteúdo pode ser acessado e alterado por qualquer usuário, bem como pode ser utilizado em trabalhos acadêmicos, livros, sites, para fins comerciais ou não, tendo a única restrição de não impor novas restrições ao conteúdo, mantendo-o livre onde quer que seja aplicado.

Construída na base do esforço voluntário, a Wikipedia possui mais de 1,6 milhões usuários que contribuem para seu crescimento em 271 línguas e dialetos³. De acordo com Jimmy Wales, “Muitos pensam na Wikipedia como 10 milhões de pessoas, cada um escrevendo uma frase. Mas na verdade, a maior parte do trabalho é feita por uma

1 Trabalho apresentado no IJ08 – Estudos Interdisciplinares da Comunicação do XVIII Congresso de Ciências da Comunicação na Região Sudeste, realizado de 3 a 5 de julho de 2013.

2 Graduanda do 9º período da Faculdade de Comunicação Social da Universidade Federal de Juiz de Fora. Bolsista do Programa de Educação Tutorial (PET), financiado pela Secretaria de Ensino Superior (SESu/ MEC) que tem como propósito integrar na graduação as atividades de ensino, pesquisa e extensão. Orientada pelo Prof. Dr. Francisco José Paoliello Pimenta. Email: fram.tava@gmail.com.

3 Estatística de maio de 2013, disponível em <<http://stats.wikimedia.org/EN/TablesWikipediansContributors.htm>>.



comunidade pequena [cerca de 1000]⁴".

O projeto teve início com Jimmy Wales, que planejava desenvolver uma enciclopédia controlada, de edição livre apenas para especialistas com Pós-Doutorado (WIKIPEDIA, 2013b). Neste período, o projeto chamava-se Nupedia. Larry Sanger passou a integrar a proposta mas o abandona quando a comunidade opta pela não diferenciação de contribuidores, sejam estes especialistas na área ou não. Neste momento, nasce a Wikipedia tal como é hoje: livre para quaisquer colaborações.

Regras da Wikipedia

Em oposição a outros projetos colaborativos (como o Linux ou o Firefox), a Wikipedia não possui um objetivo final (a criação de um software, por exemplo) nem tem um vínculo com a realidade. Um programa desenvolvido colaborativamente chega ao fim quando torna-se funcional. Já na Wikipedia, uma informação adicionada pode ser mantida mesmo que não seja funcional, verdadeira ou precisa.

Apesar de seu suporte digital, que viabiliza produção de grandes artigos (uma vez que não se tornam muito mais oneroso que outros artigos pequenos), a Wikipedia impõe aos artigos um limite de 6 – 10 mil palavras⁵, dividindo-o em múltiplas páginas quando o texto excede o valor.

Sob certa perspectiva, a Wikipedia funciona muito semelhante a uma enciclopédia, trazendo informações consolidadas de um determinado assunto, dados aceitos e de reconhecimento de um grupo de pessoas especializadas ou não. Das publicações, exclui-se apenas entradas com ensaios pessoais, resenhas críticas, textos propagandísticos, ideológicos e pesquisas originais, visando a manutenção do ponto de vista neutro em seus artigos.

Esta imparcialidade exigida está relacionada à exposição de diversos pontos de vista sobre o assunto, sugerindo que uma discussão de ideias diferentes ou opostas conceituará de forma mais justa do que tentar determinar um único posicionamento.

Para Jimmy Wales, esta é uma das políticas com melhor resultado na Wikipedia: “a maioria das pessoas é bastante razoável, até mesmo as que vêm de grupos de interesse.

4 Tradução livre da entrevista do New York Times (HAFNER, 2006).

5 O limite imposto por artigo é de até 32kb (excluindo imagens), o que resulta em 32.768 caracteres.



Eles entendem que não se pode colocar um posicionamento parcial porque ele não irá sobreviver na Wikipedia”⁶.

Emigh e Herring (2005) acreditam que essas regras tentam refletir as expectativas dos usuários da Wikipedia sobre como um texto enciclopédico deve ser, absorvendo o modelo encontrado em outras enciclopédias e tornando-se estilisticamente indistinguível destas.

Debate sobre qualidade na Wikipedia

Mesmo que se proponha como uma enciclopédia para utilização geral, a Wikipedia se diferencia em diversos pontos das publicações costumeiras, seja por sua liberdade de criação e edição, seja por sua volatilidade. Em estudos empíricos, a Wikipedia geralmente não se comporta como uma fonte usual de consulta.

Uma das principais observações feitas em relação à qualidade da Wikipedia diz respeito à temática dos verbetes, apresentando artigos variados que não são encontrados em enciclopédias tradicionais, tais como biografias de celebridades, resumos de filmes e livros, datas e anos que somam milhares de páginas com informações de curiosidades e outros tópicos considerados irrelevantes para serem incorporados em enciclopédias normais, o que reflete um interesse maior pelos temas populares do que pelos acadêmicos (ROSENZWIEG, 2006).

Para Greenstein & Devereux (2006), a Wikipedia é tudo aquilo que seus contribuidores querem que ela seja. Na Wikipedia Lusófona, por exemplo, foram encontrados 31 artigos destacados⁷ sob o tema Entretenimento (entre jogos, programas de televisão e personagens fictícios), ao passo que sob Medicina, foram encontrados 5 artigos.

As principais discussões sobre o conteúdo da Wikipedia dizem respeito à qualidade de seus textos, questionando se são precisos, completos e compreensíveis. Por se tratar de um ambiente aberto, a Wikipedia não possui mecanismos de controle de qualidade (FALLIS, 2008). Outros autores advertem sobre a consistência do conteúdo (DENNING et al, 2005; KEEN, 2007) e recomendam que a Wikipedia não seja usada

6 Tradução livre da entrevista da NPR.

7 Artigos destacados (em inglês *Features Articles*) são “artigos que a comunidade da Wikipedia identificou e avaliou como artigos de excelente qualidade” (WIKIPEDIA, 2013). Na Wikipedia Lusófona existem 555 artigos destacados; aproximadamente 1 em 1409 artigos recebem essa classificação.



como fonte oficial de informação. O Departamento de História do Middlebury College, por exemplo, proibiu seus alunos de citar a Wikipedia em trabalhos acadêmicos (COHEN, 2007).

Em 2005, a revista Inglesa Nature realizou um estudo sobre a qualidade dos artigos científicos da Wikipedia, apresentando-os juntamente com textos da Enciclopédia Britânica a especialistas, sem informá-los a qual enciclopédia pertenciam (GILES, 2005). Como resultado, foram encontrados um número igual de erros significativos em ambas e um número de erros pequenos 25% maior na Wikipedia, concluindo que a vantagem da Enciclopédia Britânica “não é tão maior, pelo menos no que diz respeito a artigos científicos”.

Em um estudo comparativo, o historiólogo Roy Rosenzweig (2006) procura avaliar o conteúdo historiográfico da Wikipedia, utilizando outras duas fontes enciclopédicas (uma especializada e outra de interesses gerais⁸). Seu estudo conclui que a Wikipedia possuía na época apenas metade dos artigos encontrados na enciclopédia especializada, no entanto, uma quantidade quatro vezes maior do que as entradas na enciclopédia de interesses gerais. O autor destaca também o conteúdo dos artigos que, diferentemente das enciclopédias tradicionais, não estavam relacionados à sua relevância histórica, mas sim à popularidade do assunto.

Apesar disso, a pesquisa conclui que a Wikipedia é em geral precisa nas informações, ainda que apresente pequenos erros nos artigos (em cerca de 16% do conteúdo analisado pelo autor), como é passível de ocorrer em outras enciclopédias. Até o momento da pesquisa, a Wikipedia mantinha uma página intitulada “Erros na Enciclopédia Britânica que estavam corretos da Wikipedia”⁹.

Outros autores abordaram a questão da confiabilidade do conteúdo da Wikipedia. Copestake (2006) centra a discussão em como um leitor decide confiar numa informação específica da Wikipedia “quando ele não pode confiar no artigo como um todo”.

⁸Rosenzweig utilizou a American National Biography Online e Encarta (enciclopédia online da Microsoft) em seu estudo, que analisou o conteúdo de bibliografias na Wikipedia. Rosenzweig pesquisou por 52 personalidades em ambas as enciclopédias.

⁹No inglês “Errors in the Encyclopedia Britannica That Have Been Corrected in Wikipedia”, disponível em <http://en.wikipedia.org/wiki/Wikipedia:Errors_in_the_Encyclop%C3%A9dia_Britannica_that_have_been_corrected_in_Wikipedia>



O fato de não ser escrita por autores especializados foi exaustivamente apontado na mídia (RICKMAN & ROBERT, 2006; STROSS, 2006) e abordado em artigos (FALLIS, 2008; CROSS, 2006). De fato, especialistas não se diferenciam de outros contribuidores na Wikipedia, o que pode gerar desinteresse por parte de especialistas, principalmente quando há um sentimento de posse sobre o artigo, já que não é objetivo da plataforma definir as autorias dos artigos¹⁰ (FALLIS, 2008).

Alfaro (2006) sugere implantações para análise da reputação do autor a fim de classificar a qualidade os artigos.

São diversos os trabalhos que se dispõem a analisar a qualidade dos artigos na Wikipedia: Blumenstock (2008) propõe a contagem de palavras como uma medida de qualidade. Cross (2006) sugere a idade do texto, Lih (2004) oferece uma análise sobre a influência do número de editores e edições na qualidade do artigo e Stivilia et al (2005) propõem 19 medidas para analisar a qualidade.

Santos e Prates (2010) analisam a maneira como a Wikipedia informam a qualidade de seus artigos, chegando à conclusão de que “a comunicação destas estratégias para os usuários apresentam problemas que podem levar a rupturas de comunicação do usuário com o sistema”. Mesmo havendo diversas estratégias que podem ser utilizadas para comunicar os usuários, as autoras encontraram recorrências em falhas de três diferentes tipos: “falta redundância de informações, uso precário dos signos estáticos/dinâmicos e excesso de complexidade das estratégias”.

No entanto, podem ser encontradas algumas ferramentas desenvolvidas por usuários para informar a qualidade do artigo. Na Wikipedia em inglês há a ferramenta *metadata*, que expõem a qualidade do artigo de acordo com as análises feitas pelos leitores e editores na página de discussão. A ferramenta indica a qualidade no título do artigo, que é apresentado numa das cores da legenda, que categoriza nove níveis de qualidade¹¹.

Um ponto crucial para Rosenzweig (2006) diz respeito à qualidade do texto, à riqueza e precisão na escolha de palavras nas enciclopédias tradicionais escritas por autores especializados. De acordo com o autor, a Wikipedia, por outro lado, apresenta textos ora

¹⁰ Apesar desse costume da comunidade de colaboradores, o estudo de Forte & Bruckman (XXXX) aponta que a tecnologia proporciona maneiras indiretas de estabelecer posse, como o histórico de edição, no qual é fácil distinguir o usuário que iniciou o verbete ou então quem realizou as maiores contribuições. O estudo, que procura compreender por que usuário contribuem com a Wikipedia, conclui que o reconhecimento de outros usuários é um fator importante para a manutenção da Wikipedia.

¹¹ <http://en.wikipedia.org/wiki/User:Pyrospirit/metadata>



“verborrágicos e monótonos” ora agitados. Esta inconsistência resulta da união de parágrafos de diferentes autores, sem haver um editor especializado para concatenar as idéias (FALLIS, 2008). De fato, a pesquisa de Elia (2006), que compara linguisticamente a Wikipedia com a Enciclopédia Britânica sugere que, apesar da Wiki apresentar maior média de palavras por post, a Enciclopédia oferecia mais densidade lexical e formalidade.

No entanto, como sugere Meyer (2006), essas comparações com enciclopédias pagas não parece ser as mais convenientes: dever-se-ia confrontar o conteúdo da Wikipedia com o das outras fontes da web que os usuários consultariam caso a Wiki não estivesse disponível. Para o autor, esta é a única maneira de apontar se é epistemologicamente favorável utilizar a Wikipedia como fonte de informações.

Não é raro o debate a respeito da instabilidade das informações na Wikipedia (COPESTAKE, 2006; NOFP, 2005), que geralmente apontam visões negativas sobre a facilidade de edição das publicações. Seu estado de beta eterno¹² possibilita alterações constantes da página. Tão fácil quanto incluir informações e atualizar conteúdos é realizar vandalismos na Wikipedia.

O blog conservador *No Oil for Pacifists* apresenta este fator como o principal problema da Wikipedia:

“como a maioria dos bens ‘livres’, não é de todo bom. Já que qualquer um pode escrever ou editar um artigo, alguns escritores serão vadios, *trolls*¹³ ou crianças, cujas contribuições diminuem, ao invés de adicionar conhecimento” (NOFP, 2005).

As ações de vandalismos na Wikipedia ocorrem quando informações falsas são incluídas deliberadamente (KITUR, 2007), assim como conteúdo ofensivo e a criação de textos propagandísticos. Greenstein & Devereux (2009) descrevem casos em que assessores de políticos norte-americanos foram descobertos alterando informações negativas de seus candidatos na Wikipedia.

Além dos atos de vandalismo, existem os *Hoaxes* (ou farsas), informações falsas passadas por verdadeiras criadas deliberadamente por usuários. A Wikipedia mantém

¹² Conceito atrelado à Web 2.0, que indica que o site (ou blog, sistema) encontra-se funcional, mas passível de modificações caso seja necessário. Sob este conceito, um produto nunca está finalizado, está sempre mudando em algum lugar.

¹³ Na internet (principalmente em ambientes como fóruns, chats, blogs ou wikis), a designação de *troll* é dada a usuários que se comportam de modo a frustrar e provocar outros usuários, desestabilizando uma discussão através de ofensas gratuitas e contradições do senso comum, rompendo com as etiquetas da internet.



uma lista de Hoaxes encontrados na plataforma mais de um ano após sua publicação (um artigo sobre um suposto assassino de Júlio César teria esteve publicado por mais de oito anos antes de ser corrigido)¹⁴.

Para Cunningham (2001), criador da plataforma Wiki, desde que o número de usuários contribuidores supere a quantidade de infratores, o bom conteúdo irá predominar sobre o vandalismo e Lih (2004) fala da assimetria que garante o balanço positivo de artigos de qualidade, pois o tempo para realizar o vandalismo é, muitas vezes, maior que o tempo para reverter a ação.

Outros autores, por outro lado, vêm a volatilidade das informações como um ponto positivo, pois desta forma a Wikipedia pode ser atualizada instantaneamente (como ocorre quando uma celebridade morre, cujo perfil é prontamente modificado). Esta característica confere uma qualidade de *auto-cura* (ROSENZWIEG, 2006), remediando rapidamente os erros descobertos.

Conclusões

Não se encerra aqui o esforço de reunir definições e conceitos sobre a Wikipedia. Trata-se apenas de um trabalho inicial na tentativa de confrontar e sintetizar a variedade de artigos e trabalhos a respeito da enciclopédia livre.

A princípio controversa, nos doze anos de sua existência, a Wikipedia conquistou espaço não apenas nas revistas científicas, mas no dia a dia dos milhões de usuários que a utilizam como fonte primária de referência.

De fato, reside neste fator a tentativa contínua de se estimar a qualidade das informações encontradas na Wikipedia – na procura de se estabelecer uma opinião precisa a respeito dos artigos escritos e editados por milhares de mãos –, se são corretas, confiáveis, verificáveis, enfim, úteis para seus usuários.

Visto seu perfil mutável e em constante atualização, não é possível, por ora, impor uma qualificação final à Wikipedia, sendo preferível nos mantermos no insonso meio-termo a impor uma definição que breve não será verdadeira.

¹⁴ *List of Hoaxes on Wikipedia*, disponível em
<http://en.wikipedia.org/wiki/Wikipedia:List_of_hoaxes_on_Wikipedia>.



Referências

- ADLER, B. T. et al. Assigning trust to Wikipedia content. In: 4th International Symposium on Wikis, 2008. **Proceedings...** ACM, 2008.
- ALFARO, L. **A Content-Driven Reputation System for the Wikipedia**, Santa Cruz: Selected Works, 2006.
- BOWMAN, Shayne; WILLIS, Chris. **We media**: How audiences are shaping the future of news and information, 2003. Disponível em:
<http://www.hypergene.net/wemedia/download/we_media.pdf>. Acesso em 3 mar. 2013.
- COHEN, N. History Department Bans Citing Wikipedia as a Research Source. **The New York Times**, 21 fev. 2007. Education. Disponível em
<<http://www.nytimes.com/2007/02/21/education/21wikipedia.html>>. Acesso em 07/05/2013.
- COPESTAKE, A. Errors in wikis: new challenges and new opportunities. In: **New text**: Wikis and blogs and other dynamic text sources. Association for Computational Linguistics: Trento, 2006.
- CROSS, T. **Puppy smoothies**: Improving the reliability of open, collaborative wikis. *First Monday*, 11 set. 2006.
- CUNNINGHAM, W. & LEUF, B. **The Wiki Way**. Quick Collaboration on the Web. Addison: Wesley, 2001.
- ELIA, A. An analysis of Wikipedia digital writing. In: **New text**: Wikis and blogs and other dynamic text sources. Association for Computational Linguistics: Trento, 2006.
- FORTE, A. & BRUCKMAN, A. Why do people Write for Wikipedia? Incentives to Contribute to Open-Content Publishing. 2005. Disponível em:
<<http://jellis.org/work/group2005/papers/forteBruckmanIncentivesGroup.pdf>>. Acesso em 13/05/2013.
- GILES, Jim. **Internet encyclopaedias go head to head**. *Nature* 438, no. 7070 (2005).
- GREENSTEIN, S. & DEVEREUX, M. Wikipedia in the spotlight. **Kellogg Case**. Kellogg School of Management, Northwestern University, 2009. Disponível em
<<http://www.kellogg.northwestern.edu/faculty/greenstein/images/htm/Research/Cases/Wikipedia.pdf>>. Acesso em 13/05/2013.
- GORMAN, G. E. A tale of information ethics and encyclopedias; or, is Wikipedia just another internet scam? **Online Information Review**, 31, 273-276. 2007.
- HAFNER, K. Growing Wikipedia Revises Its “Anyone Can Edit” Policy. **The New York Times**, 17 jun. 2006. Technology. Disponível em
<<http://www.nytimes.com/2006/06/17/technology/17wiki.html>>. Acesso em 07/05/2013.



INDIANA UNIVERSITY. “What is a troll?”. **Indiana University Knowledge Base**. Disponível em <<http://kb.iu.edu/data/afhc.html>>. Acesso em 10/05/2013.

KITTUR, Aniket et al. Power of the Few vs. Wisdom of the Crowd: Wikipedia and the Rise of the Bourgeoisie. In: **CHI**, São José, 2007.

LIH, Andrew. **Wikipedia as participatory journalism: Reliable sources?** metrics for evaluating collaborative media as a news resource. Disponível em <<http://jmsc.hku.hk/faculty/alih/publications/utaustin-2004-wikipedia-rc2.pdf>>. Acesso em: 3 maio 2013.

MEYER, B. **Defense and illustration of Wikipedia**. Retrieved August, 1, 2006.

NOFP. Open Source = Closed Minds, 2005. **No Oil for Pacifists Blog**. Disponível em <<http://nooilforpacifists.blogspot.com.br/2005/06/open-source-closed-minds.html>>. Acesso em: 07/05/2013.

NPR. **Talk of the Nation**, “Analysis: Wikipedia, Open Source, and the Future of the Web.”

SANTOS, Raquel Lara & PRATES, Raquel Oliveira. Investigando a comunicação sobre qualidade de artigos na Wikipedia para seus usuários In: **Proceedings... IX Symposium on Human Factors in Computing Systems**, pp. 271-272. Brazilian Computer Society, 2010.

STIVILIA, B.; TWIDALE, M. B.; SMITH, L. C. & GASSER, L. Assessing information quality of a community-based encyclopedia. **Proc. ICIQ**, pages 442–454, 2005.

WIKIPEDIA. **Artigos Destacados**. 2013. Disponível em <http://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Artigos_destacados>. Acesso em 10/05/2013.

WIKIPEDIA. **Nupedia**. 2013. Disponível em <<http://en.wikipedia.org/wiki/Nupedia>>. Acesso em 10/05/2013.