Wikipedia Verification Check: A Chrome Browser Extension*

Reed H. Harder
Dartmouth College
Thayer School of Engineering
Hanover, NH 03755 USA
reed.h.harder@gr.dartmouth.edu

Alfredo Velasco
Duke University
Department of Computer Science
Durham, NC 27708-0129 USA
alfredo.velasco@duke.edu

Michael Evans
Dartmouth College
Neukom Institute for Computational Science
Hanover, NH 03755 USA
michael.s.evans@dartmouth.edu

Chuankai An
Dartmouth College
Department of Computer Science
Hanover, NH 03755 USA
chuankai@cs.dartmouth.edu

Dan Rockmore†
Dartmouth College
Department of Computer Science
Hanover, NH 03755 USA
rockmore@cs.dartmouth.edu

ABSTRACT

In this paper we present Version 1.0 of an implementation of a Wiki reference parser with a light-weight plugin in the form of a Google Chrome [Google 2017] Extension with Javascript. The output of the parser is a "verification score" for any Wikipedia page, constructed from a combination of scores derived from reference accessibility and quality. The extension presented herein works from a pre-stored database of DOI and reference checks. Future versions working in real-time are also discussed. This work suggests generalizations to real-time verification checkers for arbitrary webpages and possibly even a real-time fact-checker for news platforms.

Keywords

Wikipedia; sources; references; verifiability; measurement; **ACM Classification Keywords:** H.3.4. Information Systems: Systems and Software Information Networks; H.5.3 Information Systems: Groups and Organization Interfaces – Computer-supported cooperative work

1. INTRODUCTION

With the rise of widely-available networked communication, individual and social life increasingly relies on online sources of information [Castells 1996]. Potential medical patients self-diagnose by searching online medical databases. Teachers and students use online reference material to provide education in mathematics, science, history, and art. Scientific researchers build on published work and online datasets to create new knowledge and advance our understanding of the world.

The quality of online information is therefore an important public and scholarly concern. The attentions and revelations around "fake news" that arose during the 2016 Presidential election are but the most dramatic of these. In the scholarly literature this concern is reflected in the examination of information quality from a variety of system and user perspectives. Studies examine, for example, how consumers seek health information on the Internet [Cline and Haynes 2001], whether online drug information is accurate [Clauson et al. 2008], whether Internet news sites reinforce existing political beliefs [Garrett 2009], and how scientific misinformation persists online [Kata 2010].

In this paper we focus on the notion of "verifiability" or the extent to which information can be checked for reliability, truth content, or accuracy and introduce the idea of a verifiability score that would be automatically displayed at the top of a Wikipedia page, in this case one viewed through a Chrome browser. Wikipedia is a free online encyclopedia and one of the most frequently consulted websites in the world [Alexa 2015] and we borrow the term "verifiability" from it. As a collaboratively written and edited encyclopedia with more than 24 million contributors worldwide, Wikipedia is also a complex sociotechnical system and "knowledge instrument" that relies on a variety of highly structured policies and voluntary enforcement to preserve the integrity of knowledge being conveyed [Wikipedia 2015d, Niederer and van Dijck 2010].

Our method evaluates article verifiability based on an automated analysis of references found in the article and thus extends related work on the quality and content of Wikipedia article references. At a much smaller scale, a study of a random sample of 50 country history articles

*An earlier unpublished version containing some of the material of this paper can be found on the Arxiv at https://arxiv.org/abs/1509.05631

†Contact author.

WWW'17 Companion, April 3–7, 2017, Perth, Australia.
ACM 978-1-4503-4914-7/17/04.
http://dx.doi.org/10.1145/3041021.3053364

from Wikipedia found that articles tended to refer to online sources, and disproportionately relied on news media and government websites [Luyt and Tan 2010]. In a larger scale project, a study of reference editing activity in a sample of 137,104 articles found that more mature articles are more likely to have more extensive references [Chen and Roth 2012]. And at a much larger scale, a study of 11 million citations in Wikipedia found that US sources are most common, that Google, media companies such as the New York Times, and databases such as IMDb and Census.gov dominate citations, and that primary sources are among the most persistent (and therefore "most valued") in Wikipedia articles [Ford et al. 2013]. We build on this work by providing a tool for automating the evaluation of reference quality that can be deployed on a large scale, and that can quickly articulate the prevalence of various reference quality failure modes.

While we work here in the specific and highly structured information context of Wikipedia, we believe that the basic idea would be generally useful as a browser extension, scoring verifiability for arbitrary webpages and online information sources.

Technical and practical verifiability

Though Wikipedia is much larger and more extensive than many online information sources, it provides an illustrative example of the challenges to quality that many online information sources face. At the heart of Wikipedia's collaborative processes are the "core content policies" of "verifiability," "no original research" and "neutral point of view." The most basic of these is verifiability. In Wikipedia, verifiability is the foundation of reliable knowledge. According to Wikipedia policy documents, "[a]ll material in Wikipedia mainspace, including everything in articles, lists and captions, must be verifiable." For policy purposes, verifiability "means that people reading and editing the encyclopedia can check that the information comes from a reliable source" [Wikipedia 2015c].

As with many online information sources, the most obvious challenge to verifiability in Wikipedia is a lack of citations and references. Without any reference material, it is difficult to verify whether information is true, accurate, and reliable. Thousands of words of Wikipedia policy documentation address the maintenance of verifiability through the correct use of references and citations to reliable sources. Current instructions focus on how to identify when citations are missing, how to provide those citations in such cases, and how to determine whether provided citations meet the Wikipedia standards for reliability [Wikipedia 2015a].

But it is important to note that simply providing citations and references does not automatically guarantee verifiability. Whether or not provided references and citations are accessible is less often considered as a challenge to verifiability. But it is just as important as providing the reference or citation in the first place. There are many ways that an online information source might provide citations and references and still be difficult to verify. These possible challenges fall into two analytical categories: "technical verifiability" and "practical verifiability."

"Technical verifiability" is the extent to which a reference provides supporting information that permits automated technical validation of the existence of the referenced material, based on existing technical standards or conventions. For example, books can be located with an International Standard Book Number (ISBN) or a Google Books ID, and journal articles can be located with a Digital Object Identifier (DOI). A missing ISBN or DOI certainly makes it more difficult to locate a book or article. But a provided ISBN or DOI could also be invalid or even entirely fictional, rendering the reference useless for verifying the information it supports. Thus, a Wikipedia article, all of whose book and journal references were invalid, would not be "technically verifiable" under this definition. Note that technical verifiability thus does not speak to the usefulness or relevance of the referenced material, just its existence. In particular, if all the ISBNs and DOIs corresponded to existing materials, but were mistakenly attached to the article, the article would still be perfectly 'technically verifiable,' although upon deeper inspection, clearly failing by some other measure.

"Practical verifiability," by contrast, is the extent to which referenced material is accessible to someone encountering the reference. For example, if a DOI is present but refers to a paywalled journal article, then the information it supports is practically unverifiable to someone without the additional means to access the supporting journal article. Similarly, if an ISBN is present but refers to a book that only has one extant copy in a library thousands of miles away, then the information it supports is practically unverifiable to someone without the additional means to access the supporting book.

2. OUR WIKIPEDIA SANDBOX

Wikipedia makes regular data dumps of its content available for download. We extracted 22,843,288 citations from the 3,437,650 citation-containing articles in the English Wikipedia data dump made on July 7th, 2014. Wikipedia keeps a data dump of the number of visits each article receives per hour [Wikipedia 2015b]. We aggregated the page views for each hour of the entire year and took the top 5,000 most viewed (as of July 2014) whose titles were found among the 3,437,650 citation-containing articles in the English Wikipedia.

The article sampling strategy reflected two analytical objectives. First, we wanted the sample to contain actively viewed articles rather than unmaintained or idle articles that were unlikely to motivate maintenance activity. Second, we wanted the sample to contain a range of articles in terms of official quality, rather than only focusing on the best ("featured") articles in Wikipedia. Some top articles are featured articles of enduring interest, but many are low-quality ("stub") articles that cover subjects of fleeting popularity.

Wikipedia does not strictly enforce a particular format for citations [Wikipedia 2015a]. However, several commonly used markup methods account for the majority of references in articles. Inline citations, corresponding to specific lines of text in the article, are usually formed using the "<ref>" tag in Wikipedia markup, which contains additional information about the source, often including reference type (book, journal, etc.), link if available, and other document identifiers. Citations can also appear that are not anchored to any particular piece of text: we refer to these as "free" citations. Free citations usually are marked with one of several common citation templates.

Our citation extraction pulled both inline citations and free citations from articles. Citations were categorized by citation type, either book or journal. Book citations were

checked for the presence of ISBNs or other identifying information. Journal citations were checked for DOIs and other numerical identifications.

Technical Verifiability.

ISBN numbers are the standard publishing industry identifier for books. ISBN numbers can be checked numerically for validity using check-digit algorithms for either their 10 or 13 digit versions [Hahn 2015]. ISBNs found with Wikipedia citations in the 'book' reference type specified in the Wikipedia markup were tested according to these algorithms. Out of 37,269 book citations, 29,736 book citations (79.8%) had valid ISBNs, while 3,145 (8.4%) of book citations had invalid ISBNs, and 4,388 book citations (11.8%) contained no ISBN information.

An alternative standardized book identifier is a Google Books ID. Google Books IDs were extracted from references containing Google Books links. This process did not rely on the 'book' reference type being indicated in Wikipedia markup, as this markup is inconsistent across references. Links were tested for validity using bulk submissions to a Google developer API designed for Google Books [Google 2015]. Out of 14,081 Google Books-containing citations, 3,159 (22.4%) contained invalid Google Books IDs.

Adding the presence of valid Google Books IDs as a marker of technical verifiability even in the absence of a valid ISBN, we get a slight improvement in the overall technical verifiability of book citations: 31,578 (84.7%) contain valid identifiers, 3,218 (8.6%) lack valid identifiers, and 2,473 (6.6%) contain invalid identifiers. Adding in consideration of Google Books links in other citations (not explicitly labeled "book"), we see similar proportions: 34,231 (84.7%) out of 40,381 contain valid identifiers, 3,218 (8.0%) lack valid identifiers, and 2,932 (7.3%) contain invalid identifiers.

Journal article citations were slightly more difficult to test for validity in bulk form. Instead, presence or absence of a Digital Object Identifier (DOI) was noted for any reference tagged as 'journal', 'study', 'dissertation', 'paper', 'document', or similar. Out of 41,244 of these citations, only 5,337 (12.9%) contained neither a DOI or a link to a known open access journal.

Practical Verifiability.

Verifying the open access nature of a journal citation beyond the simple presence or absence of a digital identifier is often difficult. Only a few journals are exclusively open access, and journal reference pages often have idiosyncratic layouts, making bulk web scraping for open-access confirmation challenging. Journal citations linking to 'arXiv' and 'PubMed Central (PMC)' were taken to be open access, while all others were marked unconfirmed. 5,275 of the journal citations out of 41,244 (12.8%) belonged to this confirmed open access category, while 30,632 or 74.3% contained some digital identifier but were not confirmed to be open.

Google's API allowed us to classify the accessibility of the linked Google Books into three categories: fully viewable, with all pages accessible; partially viewable, with a sample available; or not viewable at all. Out of the 10,922 working Google Books links, most (7,749, or 71.0%) are partially viewable with samples, while only 1,359 (12.4%) are fully viewable and 1,814 (16.6%) are not viewable at all.

3. VERIFIABILITY METRICS

In order to formulate and test different metrics for the verifiability of Wikipedia articles, we took proportions from the technical and practical verifiability measures calculated above, and took a weighted sum to produce an aggregate score for each page. For measures of technical validity, we looked at the proportion of valid ISBNs, and the proportion of functional Google Books identifiers. For measures of practical verifiability, we looked at the proportion of journals verifiably open access (in arXiv and PMC), the proportion of linked Google Books with fully open access, and the proportion of linked Google Books with partial access. We also considered presence or absence of numerical identifiers: the proportion of journals with a DOI, and the proportion of book citations with some sort of numerical identification (either from Google or an ISBN).

Using these measures, we constructed 4 different models of aggregate scoring, each weighting different proportions more or less heavily. Our baseline model (referred to as Model 1) weighted the technical and practical aspects of verifiability equally (with partial Google Books access conferring half the weight of a full Google Books access). So, for example, in our baseline model the article "arbitration" received a score of 2.07, "Bugatti" received a score of 3, and "Nero" received a score of 2.27. Table 1 displays the score breakdown under Model 1. Of course any one of these metrics is arbitrary and our aim in constructing and showing multiple models with different is to examine the consistency of a single aggregate metric of verifiability under varying assignments of component importance.

Table 2 reports the weighting scheme for each model. For Model 2, we weighted technical measures of verifiability more heavily. Model 3 instead weighted practical elements more heavily. Finally, Model 4 used baseline weighting for technical and practical elements, and added the two identifier categories, to reward the presence of electronic identification numbers. For each model the weighting changed the possible score for each article. For example, "Bugatti" scored a 3 in the baseline model, 5 in model 2, 4 in model 3, and 3.33 in model 4.

Article scores are only directly comparable within a model, so we ranked articles according to their individual scores under each model to get a sense of inter-model consistency, and then compared rank across models. This can be visualized as a scatter plot, with the x-axis representing articles 1 to 5,000 in descending order of score according to Model 1. Each article's corresponding rank in the model being compared is then plotted on the y-axis. As Figs. 1 and 2 illustrate, Models 2 and 3 show relative consistency in ranking with Model 1. By contrast, as Fig. 3 illustrates, Model 4 (with added identifier rankings) shows some significant variability in ranking. Block-like structures in the plot arise from regions of uniform scoring according to Model 1. Overall, we find a general correlation between the rank orderings of these different models, but with some substantive differences in individual article ranking.

To get a sense of the factors underlying divergences in ranking between models, some specific examples are illustrative. The largest gain in rank from Model 1 to Model 2 was the article "Arbitration," which gained 2,294 spots, from having a score ranked 3,931 to a score ranked 1,637. This gain makes sense in light of Model 2's emphasis on citation validity, as both of this article's ISBNs and both of its

Table 1: Examples of score calculation, Baseline model (Model 1).

Score component	(Weight)	Arbitration	Bugatti	Nero
ISBNs valid	1	1	1	1
Google Books links valid	1	1	1	1
journals with DOI	0	0	0	0
books with identifier	0	0	0	0
journals verified open access	1	0.07	0	0.07
Google Books with full/public domain access	1	0	1	0
Google Books with partial access	0.5	0	0	0.2
Article Score		2.07	3	2.27

Table 2: Weighted components for each model.

Proportion of	Model 1	Model 2	Model 3	Model 4
ISBNs valid	1	2	1	1
Google Books links valid	1	2	1	1
journals with DOI	0	0	0	1
books with identifier	0	0	0	1
journals verified open access	1	1	2	1
Google Books with full/public domain access	1	1	2	1
Google Books with partial access	0.5	0.5	1	0.5

Google Books IDs were valid. The greatest loss in rank was by the article "Microwave," which dropped 3,305 spots from rank 741 to 4,046. One of its two ISBNs was invalid, and one of its three Google Books links was broken. In short, if there are only a few references to verify, variation in weighting can give dramatic effects.

Comparing Model 1 and Model 3, the greatest gain in article rank was a 3,318 spot jump by "Glycerol" from rank 3,891 to rank 573. This article's only ISBN was invalid, explaining a low ranking under Model 1, but its one Google Books ID was fully viewable, raising the article's relative score under Model 3's emphasis on practical verifiability. The greatest drop in this comparison was the article "Nero," which dropped 1,903 places from rank 1,632 to 3,535, hurt under greater emphasis on practical verifiability with three out of its five Google Books IDs being completely unavailable for free online viewing.

Comparing Model 1 and Model 4, the greatest gain in rank was by "Pneumothorax," which jumped 2,497 places from rank 3,856 to rank 1,359. With Model 4's added weighting for the presence of identifiers, this article was helped by the fact that all 24 of its journals had electronic identification (DOI, or confirmed open access), and seven out of its nine book links contained either a valid ISBN or Google Books ID. The greatest drop in rank under Model 4 was by "Bugatti," which dropped 3,931 places from rank 74 to rank 4,005. Both of its journal citations had no electronic identification and two out of its three books contained neither an ISBN nor Google Books link.

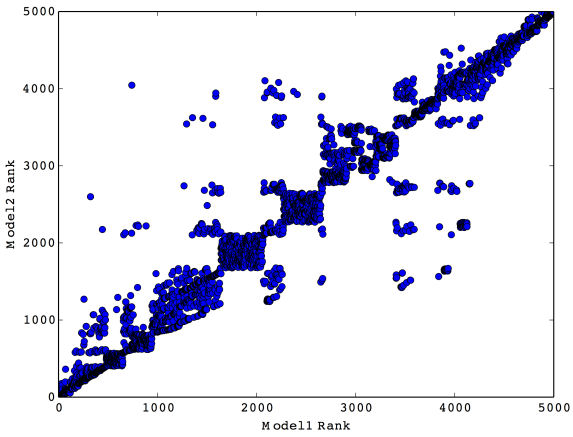

Figure 1: Change in article verifiability rank, baseline model (Model 1) vs. Model 2.

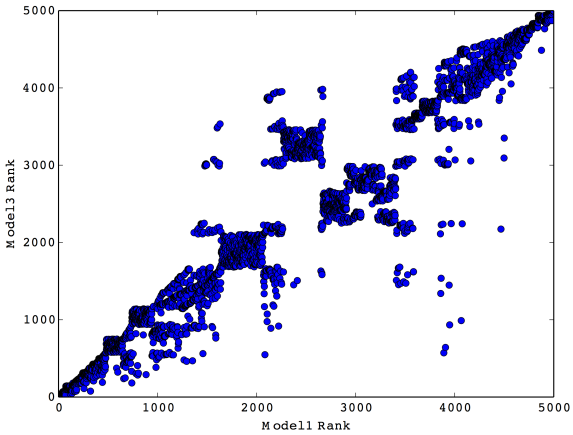

Figure 2: Change in article verifiability rank, baseline model (Model 1) vs. Model 3.

4. CHROME EXTENSION

The preceding sections detail the kinds of considerations that have gone into our choice of verifiability metric or verifiability score. In this section we briefly discuss an extension to the Chrome browser that when presented with a Wikipedia page would then display the verification score as well as some of the information that has gone into its computation. As discussed our current version works only on

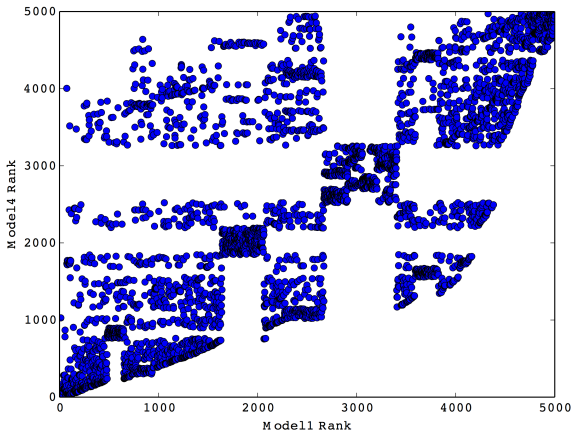

Figure 3: Change in article verifiability rank, baseline model (Model 1) vs. Model 4.

the sandboxed Wikipedia dataset of 5,000 pages described above. Once the Wiki extension has been loaded to Chrome, when the current tab visits a Wikipedia page and the user clicks the icon of the Wiki extension, it will read the title and URL of the Wikipedia page and query the Verification Score from a local database. The verification score has been computed and stored in a database. The database returns the Verification Score with some statistics and details about the reference links on the current Wikipedia page. Finally, the Wiki extension displays the query results in a pop-up textbox. Figure 4 is an example of the user interface for the current Wikipedia page "Javascript".

4.1 Real-time vs. Offline

The implementation we describe thus assumes an offline calculation of the score – depending on an offline evaluation of the various verification components. Extensions of this simple approach to the full Wikipedia dataset would require a continuous crawling of Wikipedia and updating of the metrics which would be facilitated by the necessary continuous interactions between the Wikipedia references and DOI validation, URL checking, etc.

The offline extension reads data directly from a database so it saves time during operation. However, it is useful to consider a real-time extension since many Wikipedia pages are regularly updated. A real-time Wikipedia parser/verifier should (1) parse html codes to detect external references (2) classify those references into several groups (e.g., books, ISBN) and (3) compute the verification score. The bottleneck for a real-time extension is the longer running time coming from the need to make multiple queries for the verifications. A practical approach might be to combine offline and real-time modes together, perhaps using a database to store only the Verification Scores of the top K most frequently visited Wikipedia pages. For the remainder a real time approach could be applied.

Discussion

We have presented an approach to measuring verifiability on Wikipedia pages and a simple application for deploying it as an extension to the Chrome browser.

Some caveats are worth noting that suggest interesting directions for future research around verification. Note that the models only consider obviously Open Access sources such as PubMed and arXiv, and might productively be expanded to include other sources known to be Open Access (e.g., listed in the Directory of Open Access Journals). Also, as indicated above, the measures in these simple models do not account for the small number of references in Wikipedia articles, and might productively be extended in future research to be more robust. Of particular relevance is a study of Willinsky that showed that with some work, open source references can sometimes be found to replace non-open references [Willinsky 2007].

Future research on Wikipedia might also examine in more detail the variation in verifiability scores between individual articles in different models. Using fractions makes the models more robust for articles with many references, so rankings for a single article with few book or journal references can change significantly even if the change in number or validity of references is small in absolute terms. This suggests future opportunities for considering reference density and reference quality together in the study of verifiability. One possible direction would be investigating effects of genre or category on verifiability. There may well be informal, genre-specific editorial expectations that favor one model of verifiability over another. Similarly, comparison of article verifiability rankings against Wikipedia's internal article quality rankings could provide useful insight. While previous work has noted a relationship between article age and density of references [Chen and Roth 2012], a consideration of reference quality might illuminate more complex relationships between article quality and article maturity.

A more practical direction would be to incorporate an article-level verifiability metric into the Wikipedia browsing experience, allowing users to compare the empirical reality of verifiability against broader policy expectations. Connecting verifiability to user experience would also address verifiability as a potential source of user inequality and bias in Wikipedia articles. The burden of satisfying the verifiability metric currently falls on editors who may have very different access to, and preferences for, reliable knowledge [Wikipedia 2015c]. Making verifiability visible to users could encourage wider participation by users with different perspectives on access to knowledge.

At the individual reference level, our method may be useful in the automated large-scale flagging of problematic references, or in the identification of users associated with a large number of such references. Our article-level verifiability could also be incorporated into higher-level systems addressing article quality as a whole. For example, recent work has examined the automated classification of hoax articles based on article content [Kumar et al. 2016]. It is possible that such methods could be improved by considering the reference quality and accessibility features measured here. It is possible that such hoax articles frequently cite technically functional and accessible information, but information unrelated to the claims made in the article. This issue may motivate a more detailed examination of reference verification that evaluates the similarity of the article text and the cited text, with the goal of flagging references below a certain similarity threshold. While such an approach presents both technical challenges (e.g., parsing the source text of a variety of referenced sources) and theoretical challenges

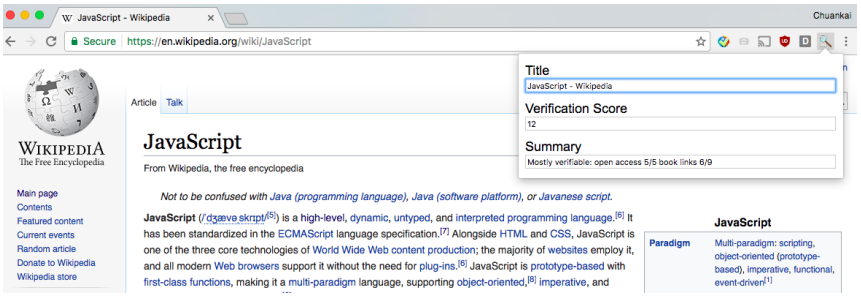

Figure 4: Example of user interface.

(e.g., evaluating the relationship of a text similarity threshold and relevance to a factual claim), it may be a highly fruitful avenue for future work.

Our approach to constructing a flexible and customizable verifiability metric helps make visible potential problems of verifiability, and increases the possibility for improving verifiability for all users in Wikipedia. But what is possible for Wikipedia is also possible for many other online information sources. For example, this approach could be extended to measure the practical verifiability of scientific papers by looking at whether their supporting citations and data are readily available for review. Similarly, this approach could be extended to a browser extension that scrapes any citation or reference on a web page and calculates that page's technical and practical verifiability. Such an approach would address several problems of verifiability around "fake news" (see e.g., [Rubin et al. 2016]) or other unsourced, unverifiable, or low information quality web pages (cf., [Conroy et al. 2015] for a survey of some approaches). In fact, the concept of a hoax detection platform has already been discussed [Shao et al. 2016]. However this approach is extended in the future, measuring verifiability will help address variations in the quality of references in online information and, ideally, improve their overall quality.

5. ACKNOWLEDGMENTS

We gratefully acknowledge the support of the Neukom Institute for Computational Science. We also acknowledge the helpful feedback received from several anonymous readers.

6. REFERENCES

[Alexa 2015] Alexa. 2015. Alexa Top 500 Global Sites. http://www.alexa.com/topsites. (2015).

[Castells 1996] Manuel Castells. 1996. *The Rise of the Network Society*. Blackwell Publishers, Cambridge, MA.

[Chen and Roth 2012] Chih-Chun Chen and Camille Roth. 2012. Citation Needed: The Dynamics of Referencing in Wikipedia. In *Proceedings of the 8th Annual International Symposium on Wikis and Open Collaboration*. ACM.

[Clauson et al. 2008] Kevin A. Clauson, Hyla H. Polen, Maged N. Kamel Boulos, and Joan H. Dzenowagis. 2008. Scope, Completeness, and Accuracy of Drug Information in Wikipedia. *Annals of Pharmacotherapy* 42, 12 (2008), 1814–1821.

[Cline and Haynes 2001] R. J. W. Cline and K. M. Haynes. 2001. Consumer Health Information Seeking on the Internet: The State of the Art. *Health Education Research* 16, 6 (2001), 671–692.

[Conroy et al. 2015] Niall J. Conroy, Victoria L. Rubin, and Yimin Chen. 2015. Automatic deception detection: Methods for finding fake news. *Proceedings of the Association for Information Science and Technology* 52, 1 (2015), 1–4. DOI:http://dx.doi.org/10.1002/pra2.2015.145052010082

[Ford et al. 2013] Heather Ford, Shilad Sen, David R. Musicant, and Nathaniel Miller. 2013. Getting to the Source: Where does Wikipedia Get its Information From?. In *Proceedings of the 9th International Symposium on Open Collaboration*. ACM, New York, 9:1–9:10.

[Garrett 2009] R. Kelly Garrett. 2009. Echo Chambers Online?: Politically Motivated Selective Exposure Among Internet News Users. *Journal of Computer-Mediated Communication* 14, 2 (2009), 265–285.

[Google 2015] Google. 2015. Google Developers: Using the API - Google Books. https://developers.google.com/books/docs/v1/using. (2015).

[Google 2017] Google. 2017. Chrome Web Browser. https://www.google.com/chrome/. (2017).

[Hahn 2015] Joel Hahn. 2015. ISBN Check Digit Calculator. http://www.hahnlibrary.net/libraries/isbncalc.html. (2015).

[Kata 2010] Anna Kata. 2010. A Postmodern Pandora's Box: Anti-Vaccination Misinformation on the Internet. *Vaccine* 28, 7 (2010), 1709–1716.

[Kumar et al. 2016] Srijan Kumar, Robert West, and Jure Leskovec. 2016. Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. In *Proceedings of the 25th International Conference on World Wide Web (WWW '16)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 591–602. DOI: http://dx.doi.org/10.1145/2872427.2883085

[Luyt and Tan 2010] Brendan Luyt and Daniel Tan. 2010. Improving Wikipedia's Credibility: References and Citations in a Sample of History Articles. *Journal of the Association for Information Science and Technology* 61, 4 (2010), 715–722.

[Niederer and van Dijck 2010] Sabine Niederer and José van Dijck. 2010. Wisdom of the Crowd or Technicity of Content? Wikipedia as a Sociotechnical System. *New Media & Society* 12, 8 (2010), 1368–1387.

[Rubin et al. 2016] Victoria L Rubin, Niall J Conroy, Yimin Chen, and Sarah Cornwell. 2016. Fake News or Truth? Using Satirical Cues to Detect Potentially Misleading News.. In *Proceedings of NAACL-HLT*. 7–17.

[Shao et al. 2016] Chengcheng Shao, Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menczer. 2016. Hoaxy: A Platform for Tracking Online Misinformation. In *Proceedings of the 25th International Conference Companion on World Wide Web (WWW '16 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, Switzerland, 745–750. DOI: http://dx.doi.org/10.1145/2872518.2890098

[Wikipedia 2015a] Wikipedia. 2015a. Wikipedia:Citing Sources. https://en.wikipedia.org/wiki/Wikipedia:Citing_sources. (2015).

[Wikipedia 2015b] Wikipedia. 2015b. Wikipedia:Index of Page View Statistics for 2014. https://dumps.wikimedia.org/other/pagecounts-raw/2014/. (2015).

[Wikipedia 2015c] Wikipedia. 2015c. Wikipedia:Verifiability. Online; accessed 23-March-2015. (2015).

[Wikipedia 2015d] Wikipedia. 2015d. Wikipedia:Wikipedians. https://en.wikipedia.org/wiki/Wikipedia:Wikipedians. (2015).

[Willinsky 2007] John Willinsky. 2007. What open access research can do for Wikipedia. *First Monday* 12, 3 (2007).