

The Use and Operationalization of “Misinformation” and “Disinformation” by Wikipedia Editors

Joseph S. Schafer

University of Washington
Human Centered Design & Engineering
schaferj@uw.edu

Mark Zachry

University of Washington
Human Centered Design & Engineering
zachry@uw.edu

Keywords: misinformation, disinformation, edit annotations, mixed-methods, editing behavior

Introduction

Misinformation and disinformation are issues of increasing societal attention, both within and outside the academy. However, while these concepts have been increasingly studied and debated, less research attention has thus far been directed towards how communities operationalize these terms. Inspired by (Farkas and Schou, 2018)’s exploration of the term “fake news,” we analyze how Wikipedia editors have used these social terms and concepts in their work. Our study contributes to Wikipedia research in two important ways. First, by sampling our data from edit annotations on article and talk pages, we use a relatively novel, under-explored method for Wikipedia research (outside of (Ekstrand and Riedl, 2009)), and demonstrate how this can reveal important data which may otherwise remain unseen. Second, this research focuses on an important deliberative practice by Wikipedia editors over one measure of article content quality, and provides empirical evidence of changes in editors’ usage of the concepts of mis- and disinformation.

Methods

To understand how Wikipedia editors operationalize the *concepts* of mis- and disinformation, we look to see how the terms “misinfo,” “disinfo,” “misinformation,” and “disinformation” are used. To differentiate between the concepts and the lexical strings, we will refer to the lexical strings in quotes throughout, while the concepts will be left outside of quotes.

We collected every edit from English-language Wikipedia in the article and article-talk namespaces from the Wikipedia data dumps, approximately 550 million edits, from its inception in 2001 until March 2023. We then filtered these edits to only those containing the substring “misinfo” or “disinfo” in their edit annotations, and removing the terms “disinfopedia,” “misinform,” “misinformed,” and “misinforming.” We then filtered out bots (using the method in (Geiger and Halfaker, 2017)), resulting in a set of 41,860 edits with these terms.

To supplement the edits dataset, the authors independently qualitatively coded (dual-coded) all articles which

had at least five edit annotations mentioning the terms in a given year in either the article namespace or five edits in a given year mentioning the terms in the talk page namespace. Additionally, since editing volume in the earliest years of the encyclopedia was significantly smaller, we dual-coded every article and talk page mentioning the terms at least once in 2003 or earlier. This resulted in a curated set of 493 articles (representing 7,563 of these edits), of which we later excluded 13 pages containing 67 edits, since those pages were no longer live.

We coded two criteria: **What** topic the article focused on, and **where** (if anywhere) the terms “misinformation” or “disinformation” appeared beyond the annotations.

To categorize the topics of the articles referenced, we started with a list of topics prevalent in mis/disinformation research, informed by the first author’s four years of working in misinformation research, and in pilot iterations of coding iteratively expanded these categories to capture categorical topics of the articles that did not fit. This eventually resulted in a list of 18 named categories, plus an “other” category. Disagreements in coding were discussed and arbitrated to consensus between the two authors. To categorize where terms appeared, the authors used text-finding tools on the article text, talk page, and talk page archive search (if available) for “misinfo” and “disinfo,” manually evaluating all matches. Note, since Wikipedia pages are living documents, these location tags are accurate to when the coding was conducted, in Autumn of 2023.

We then created time series and distribution plots, describing how the frequency of edits mentioning these terms changed over time, and how these terms are distributed both in locations across the encyclopedia and in the varying topics editors are engaging. This topical analysis builds on the second author’s previous experiences in coding social behaviors and language use on Wikipedia (McDonald and Zachry, 2021).

Results

Prevalence Over Time

First, we graph the number of edits mentioning these terms over time as a percentage of overall edits by year, as shown in Figure 1. There is an increase in edits mentioning these terms post-2015, which lines up with existing

research (Freelon and Wells, 2020). However, there is also a spike in edits mentioning these terms earlier around 2003-2006. Many of these edits come from discussions of the second U.S.-Iraq War. Additionally, approximately when this initial spike fades is around when Wikipedia established the core policy of verifiability, potentially meaning that editors coalesced some of their discussions using this concept, rather than mis/disinformation.

Locations Within Articles

Our research shows that among our 480 sampled articles, 141 had mentions of “misinfo” or “disinfo” *only* in the edit annotations (for the full distribution see Table 1). These distributions also changed over time, with notable increases in article text usage, as shown in Figure 2.

Article Topic Distribution

Our findings also illustrate significant topical diversity among the articles where editors mention mis/disinformation. This includes subject areas not commonly studied in misinformation studies, such as articles about pop culture. The distribution is shown in Figure 3.

Discussion and Conclusions

Our findings illustrate that while not primary terms used in Wikipedia edits, the language of misinformation and disinformation are nevertheless meaningfully used by editors and correspond to varied community practices. Editors adapt these concepts into their own work, both bringing these terms into newer contexts and using their own definitions of these terms for their own purposes. One notable example of this is an editor writing on the “Star of David” talk page that information is “not disinformation if it’s true,” an intuitive definition but one out of step with current academic use of the term, which encompasses both false and true-but-misleading information.

Beyond the empirical findings, this work illustrates that edit annotations are a rich additional data source within publicly available Wikipedia data. When reviewing literature for this project, and building on the second author’s experience conducting research on Wikipedia, we noted few studies used edit annotation data for primary analysis, with the notable exception of (Ekstrand and Riedl, 2009). Starting our analysis with this data stream surfaced uses of these concepts which would otherwise remain invisible, as we found 141 articles where the edit annotations were the only location where these concepts were acknowledged. This included instances where edit was referenced misinformation in their annotations as discussions between editors. We also found cases where talk page discussions referencing misinformation were incorrectly archived by bots, leaving their only active trace in the edit annotations. While we acknowledge this approach misses the converse data (where the terms appear

in-text but not in annotations), we believe this approach complements other methods which focus more on text in articles or in talk pages, collectively expanding our understanding of editor behaviors in a largely neglected space of editor engagement and interaction.

This research also outlines important future research directions, such as conducting discourse and content analyses of pages where these terms appear, describing the kinds of editors who use these concepts, understanding the use of related terms such as those cataloged in (Jack, 2017), and exploring how these terms are used in other Wikipedia areas, such as policy discussions.

Acknowledgments

We thank Kate Starbird, Soobin Cho, Pitch Sinlapanuntakul, Julie Vera, and the reviewers for their feedback, and Kaylea Champion and the Community Data Science Collective for help acquiring data. Joseph is supported by a National Science Foundation Graduate Research Fellowship, Grant DGE-2140004. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

References

- [Ekstrand and Riedl2009] Michael D. Ekstrand and John T. Riedl. 2009. rv you’re dumb: identifying discarded work in Wiki article history. In *Proceedings of the 5th International Symposium on Wikis and Open Collaboration, WikiSym ’09*, pages 1–10, New York, NY, USA, October. Association for Computing Machinery.
- [Farkas and Schou2018] Johan Farkas and Jannick Schou. 2018. Fake News as a Floating Signifier: Hegemony, Antagonism and the Politics of Falsehood. *Javnost - The Public*, 25(3):298–314, July.
- [Freelon and Wells2020] Deen Freelon and Chris Wells. 2020. Disinformation as Political Communication. *Political Communication*, 37(2):145–156, March. Publisher: Routledge eprint: <https://doi.org/10.1080/10584609.2020.1723755>.
- [Geiger and Halfaker2017] R. Stuart Geiger and Aaron Halfaker. 2017. Operationalizing Conflict and Cooperation between Automated Software Agents in Wikipedia: A Replication and Expansion of ‘Even Good Bots Fight’. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW):49:1–49:33, December.
- [Jack2017] Caroline Jack. 2017. Lexicon of Lies: Terms for Problematic Information.
- [McDonald and Zachry2021] David W. McDonald and Mark Zachry. 2021. On the Alignment Between Self-Declared Gender Identity and Topical Content from Wikipedia. *ACM Transactions on Social Computing*, 4(2):7:1–7:69, June.

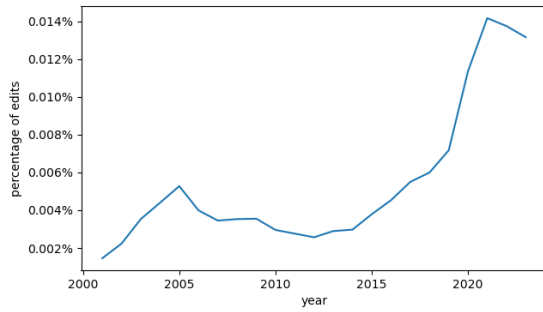


Figure 1: Percentage of edits per year mentioning “misinfo” or “disinfo” in their edit annotations.

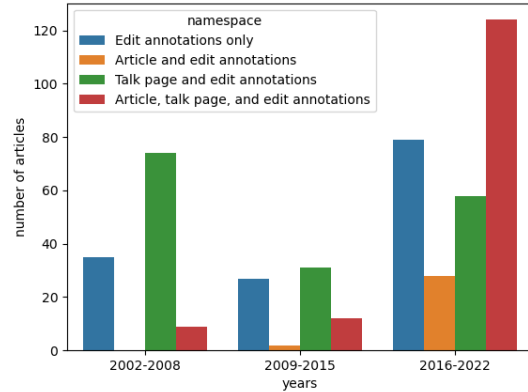


Figure 2: The distribution of article-talk page pairs over time, by which namespace the terms misinformation and disinformation appear. Note, for this figure we only focus on years for which we have complete data, omitting 2001 and 2023. In addition, we grouped years into three buckets, since the number of articles in a particular year was often quite small.

Namespace	# of Articles
Article and edit annotations	31
Article, talk page, and edit annotations	145
Edit annotations only	141
Talk page and edit annotations	163

Table 1: The number of articles where the terms misinfo or disinfo appeared in particular namespaces.

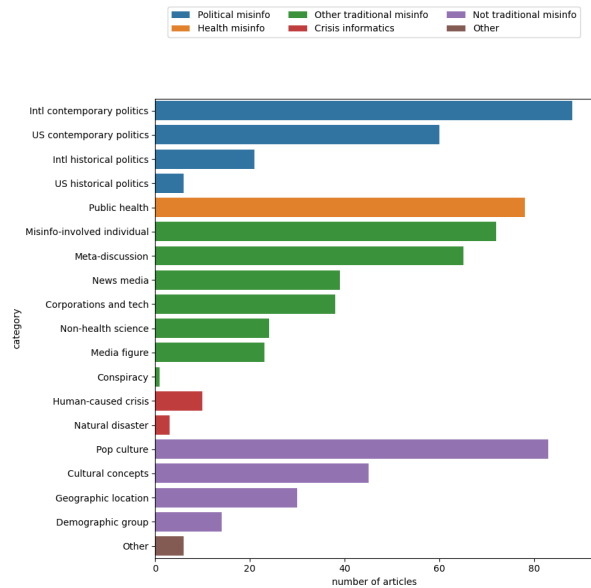


Figure 3: The distribution of articles edited with the terms misinformation and disinformation, organized by category. We also categorized the categories themselves, sorting the articles into fitting subfields of mainstream mis/disinformation research, and related research fields.