AI & SOCIETY
https://doi.org/10.1007/s00146-025-02199-9

RESEARCH

An endangered species: how LLMs threaten Wikipedia's sustainability

Matthew A. Vetter[1] · Jialei Jiang[2] · Zachary J. McDowell[3]

Received: 18 September 2024 / Accepted: 22 January 2025

Abstract
As a collaboratively edited and open-access knowledge archive, Wikipedia offers a vast dataset for training artificial intelligence (AI) applications and models, enhancing data accessibility and access to information. However, reliance on the crowd-sourced encyclopedia raises ethical issues related to data provenance, knowledge production, curation, and digital labor. Drawing on critical data studies, feminist posthumanism, and recent research at the intersection of Wikimedia and AI, this study employs problem-centered expert interviews to investigate the relationship between Wikipedia and large language models (LLMs). Key findings include the unclear role of Wikipedia in LLM training, ethical issues, and potential solutions for systemic biases and sustainability challenges. By foregrounding these concerns, this study contributes to ongoing discourses on the responsible use of AI in digital knowledge production and information management. Ultimately, this article calls for greater transparency and accountability in how big tech entities use open-access datasets like Wikipedia, advocating for collaborative frameworks prioritizing ethical considerations and equitable representation.

Keywords Large language models (LLM) · Aritificial intelligence (AI) · Wikipedia · Sustainability

1 Introduction

A collaboratively edited, open-access knowledge archive, Wikipedia, provides a vast dataset for training artificial Intelligence (AI) applications and models (Deckelmann 2023; Gertner 2023; Liu et al. 2024; McDowell 2024; Schaul et al. 2023). While such repurposing can make the encyclopedia's content more accessible, it also introduces numerous ethical issues related to data provenance, knowledge production and curation, and digital labor. Such issues, especially those related to labor and disintermediation, have been hypothesized as a threat to Wikipedia's overall sustainability (Wagner and Jiang 2025). Drawing on critical data studies (boyd and Crawford 2012; Iliadis and Russo 2016), feminist posthumanism (Haraway 1988, 1991), and recent critical interrogations of Wikidata's ethics (Ford and Iliadis 2023; McDowell and Vetter 2024; Zhang et al. 2022), this article explores the potential biases and power dynamics embedded in the data curation processes of Wikipedia and its subsequent use in large language models (LLMs). Our research employed a problem-centered expert interview (Döringer 2020) to investigate a complex, current issue: the relationship between Wikipedia and LLMs in connection with issues of sustainability, information access and literacy, problematic information, and systemic bias. While Wikipedia's open-editing model democratizes data creation, the algorithms used by LLMs to process and prioritize this data can perpetuate systemic biases, thus influencing public perception and access to information. In addition to this issue, AI-powered LLMs also threaten Wikipedia's long-term sustainability and maintenance, as they effectively detour the website and its capacity to recruit new editors. Our article calls for greater transparency and accountability in how tech giants leverage open-access datasets like Wikipedia, advocating for collaborative frameworks that prioritize ethical considerations and equitable representation. By foregrounding these concerns, this study contributes to ongoing discourses on the responsible use of AI in digital knowledge

✉ Matthew A. Vetter
mvetter@iup.edu

Jialei Jiang
j.jiang@pitt.edu

Zachary J. McDowell
zjm@uic.edu

1 Indiana University of Pennsylvania, Indiana, USA

2 University of Pittsburgh, Pittsburgh, USA

3 University of Illinois at Chicago, Chicago, USA

Published online: 19 February 2025

production and information management to address the following research questions.

1. How are large language models (LLMs) trained on Wikipedia, and what specific aspects of Wikipedia's content contribute to their development?
2. In what ways does the integration of LLMs in AI-powered chatbots such as ChatGPT, Microsoft's Copilot, and Google's Gemini impact Wikipedia's sustainability as a crowd-sourced knowledge platform?
3. What challenges related to information literacy and digital labor exploitation arise from the use of Wikipedia-sourced content in AI-powered chatbots?

In the following, we first review the relevant literature to define sustainability and theorize Wikipedia as a data archive, placing particular emphasis on how its policy and practices around verifiability ("Wikipedia:Verifiability" 2024) make it an ideal training set for LLMs. Following a discussion of our theoretical framework and research methods, we present six key findings with summary and quotations, drawn from the interview data. Finally, we place these findings in conversation with previous research and offer recommendations for relevant stakeholders.

2 Wikipedia and Sustainability

This study addresses the construct of sustainability, which has been defined as a multidimensional organizational property that integrates social, environmental, and financial considerations (among others) (Giovannoni and Fabietti 2013). The complexity of sustainability necessitates an integrated approach capable of managing interrelationships across these considerations, especially in the case of peer production (Pestoff 2014). Although Wikipedia has been celebrated in the past for the effectiveness of its governance model, which emphasizes community control and embraces diverse forms of participation (Morrell 2014), its continued relevance has been called into question more recently given the impact of generative AI on information ecosystems (Gertner 2023; Wagner and Jiang 2025).

To understand sustainability in the context of Wikipedia, we need to see it as more than just a website or platform – it is a community of people dedicated to maintaining "the essential infrastructure for free knowledge" (Wikimedia Foundation 2024). The Wikimedia Foundation, Wikipedia's nonprofit parent organization, has both defined and set goals around sustainability as part of its movement strategy recommendations. According to these recommendations, sustainability involves (1) supporting and investing in the needs of all contributors, (2) adopting equitable approaches to resource generation and distribution, and (3) recognizing diverse contributions beyond content creation. This includes public policy, advocacy, capacity building, outreach, research, organizing, and fundraising ("Movement strategy" 2024). While these recommendations are comprehensive enough to speak for the entire Wikimedia movement, they also help to elucidate sustainability in the specific context of the Wikipedia project. For Wikipedia to survive as a reliable and comprehensive encyclopedia, it must continuously attract and keep a diverse group of dedicated volunteers. These contributors work together to create and maintain high quality content while fostering and governing community policy. Sustainability also involves building systems for quality control that standardize assessment and peer review in the community (Lichtenstein and Parker 2009). Finally, a sustainable Wikipedia needs a steady stream of new and returning readers, those who have the means to contribute to the project both in terms of future potential volunteering and/or donating to the cause.

3 Wikipedia as data archive

A data archive is broadly understood as a repository that preserves and provides access to information over time. For instance, Bowker (2010) states that an archive is "the set of all events which can be recalled across time and space" (212). Although most archives are made up of "potential memory" (212), they also include imperative elements, recalling only what is needed at a particular time for a particular purpose. While Bowker (2010) focuses on a broad definition of the archive, Borgman et al. (2015) define the digital data archive more specifically. Digital data archives, as outlined by Borgman et al. (2015), are vital components of contemporary scholarly communication and knowledge infrastructures. These repositories vary widely, ranging from domain-specific databases to generic platforms.

As a digital archive, Wikipedia operates as a vast, collaborative repository of human knowledge that supports both immediate access and sustained documentation of historical and cultural information. Wikipedia's role as a digital data archive aligns with the principles of "Archives 2.0" (Cooban 2017: p. 269) which prioritize openness, user participation, and flexibility in archival practices. According to Cooban (2017), Wikipedia exemplifies "participatory archives" (p. 269) where archivists act as facilitators rather than gatekeepers; e.g., contributing to articles, linking collections, and fostering greater accessibility. Wikipedia's collaborative model and extensive metadata system, including internal links and categories, allow the encyclopedia to serve as a dynamic platform for knowledge discovery, though it does not claim authority over the knowledge it hosts. Other platforms, such as DBWiki, expand on Wikipedia's archival potential by combining wiki functionality with database features. Buneman et al. (2011) highlight DBWiki's capacity for data versioning, provenance tracking, and annotation, available through its markdown language for embedding queries. While more specialized than Wikipedia, DBWiki unveils the importance of integrating database structures with collaborative systems to enhance data archives.

Scholars further underscore Wikipedia's archival potential as a sociotechnical system that integrates human contributors and automated processes, such as bots (Fichman and Hara 2014). Its metadata-driven architecture enhances discoverability and interoperability of knowledge (Sugimoto et al. 2015), while also linking digitized archival assets to institutional repositories boosts the visibility of specific items (Szajewski 2013). This integration with institutional repositories allows Wikipedia to bridge institutional knowledge silos and reach wider audiences. Much of what makes Wikipedia so valuable as a digital data archive, of course, is its dynamism. When Wikipedia's data is used as a training set rather than an active archive, it loses the benefits of dynamic updates, collaborative curation, structured metadata, and user participation.

3.1 Wikipedia's verifiability

As a fundamental principle of Wikipedia, verifiability refers to the practice of ensuring that claims are supported by reliable sources (Petroni et al. 2023; Redi et al. 2019; Wong et al. 2021). Wong et al. (2021) highlight how the reliability and quality of Wikipedia's content are crucial, not only for human users but also for AI systems that utilize Wikipedia as training data or source of information. These scholars further emphasize that the quality of Wikipedia's content is vital because AI systems trained on it rely on accurate information for fact-checking. Redi et al. (2019), for example, created a taxonomy of reasons why citations are necessary on Wikipedia, which can be categorized based on factors such as quotations, statistics, controversial claims, and unclear sources. This taxonomy serves as a framework for understanding the automated process of fact-checking available through LLMs.

Indeed, AI models, particularly machine learning and information retrieval algorithms, have demonstrated their capacities for maintaining and improving the verifiability of Wikipedia content. Wong et al. (2021) discuss the potential Wiki-Reliability, a dataset of Wikipedia articles annotated for content reliability, used to train AI models to predict and identify content issues. Petroni et al. (2023) examine "SIDE," an AI system designed to identify unreliable citations on Wikipedia and recommend more suitable alternatives by analyzing claims and contexts as well as searching for evidence on the internet.

Despite the current technological advances of LLMs, however, these models often struggle with generating verifiable information and content. While scholars have addressed the potential of LLMs, they have simultaneously voiced concerns over the use of LLMs to improve Wikipedia's content. Citation, or indeed any documentation of sources, is largely absent in (most) current LLM models, which not only obscures data provenance but can also contribute to user misunderstanding and ethical dilemmas. Huang and Chang (2023) argue that by attributing information to its source, citations can help mitigate plagiarism, credit original authors, and allow users to verify the generated content. Without citations, users may incorrectly attribute the LLM's output as its own opinion or creation, rather than information derived from a source, which can lead to the spread of misinformation and failure to credit original authors.

Furthermore, LLMs are also limited when it comes to replicating human processes of fact-checking. While AI can assist humans with identifying potential issues and suggesting improvements, human editors remain essential for tasks that would require contextual understanding and in-depth reasoning. Researchers have highlighted the important contributions of human editors: "AI is high-recall and low-precision compared to Wikipedia editors; models generally change the text that editors change, and much more" (Ashkinaze et al. 2024: p. 12). This observation reveals the limitations of AI in understanding contextual information as well as human editors. As such, scholars further acknowledge the challenges and limitations of relying solely on LLM for citation-related tasks. While LLMs can efficiently process large datasets and identify potential citation issues, researchers caution against viewing these models as a replacement of human judgment. Ashkinaze et al. (2024) emphasize that evaluating citations often requires a deeper understanding of context, source reliability, and potential biases. In these areas, human editors are superior to current AI models. For example, while an AI might flag a citation as potentially problematic, human editors might determine that the source is reputable within a specific field. There may also be incidents where certain controversial information is accurately represented within the article's context.

3.2 Wikipedia as a training set

The massive amount of multilingual textual data in Wikipedia, covering a wide range of topics, makes it a valuable resource for training AI models. Wikipedia's content and structure are used for a variety of LLM research purposes. Srinivasan et al. (2021) have explored the creation of the Wikipedia-based Image Text (WIT) dataset, which uses Wikipedia's content to train LLMs for image-text retrieval, cross-lingual representation, and other multimodal, multilingual tasks. This dataset's large-scale and multilingual nature—"a curated set of 37.5 million entity-rich image-text examples with 11.5 million unique images across 108 Wikipedia languages" (Srinivasan et al. 2021: p. 2443)—makes it a valuable resource for advancing research in multilingual, multimodal AI. Wikipedia's structure (Thomas 2023) contains millions of articles, interlinked pages, and a collaborative editing process, which offers a unique opportunity for training LLMs to understand and navigate complex information networks. Researchers can leverage this structure to develop AI systems that are capable of identifying patterns, relationships, and hierarchies within their datasets. Ethical uses of algorithms also carry additional communal benefits (Jiang et al. 2024; Vetter et al. 2024b). For instance, researchers can develop LLMs to reduce Wikipedia's community workload (Smith et al. 2020), maintain human judgment in decision-making, support diverse workflows, foster positive engagement with editors (especially newcomers), and establish trust in both people and algorithms.

Despite the usefulness of Wikipedia as a training set for LLMs, several concerns have been raised regarding such training, including (1) biased information, (2) insufficient transparency, and (3) labor exploitation. First, there exists significant potential for replicating existing biases in Wikipedia's content due to its volunteer-driven nature and editor demographics (McDowell 2024; Petroni et al. 2023). Researchers have also called attention to biased information, such as gender, racial, linguistic, and cultural biases in Wikipedia's content (Gruwell 2015; Jiang and Vetter 2020a, 2020b; McDowell 2021; McDowell 2024; Vetter et al. 2024a, 2018). Since LLMs learn from the data they are trained on, any existing biases within Wikipedia can be carried forward into the AI models' outputs. Further complicating the issue is that technical bias mitigation methods (Crawford 2021), such as diversifying datasets or adjusting algorithms for statistical parity, may fail to tackle the underlying problem of how power structures perpetuate social inequalities. Thus, it becomes crucial for scholars and practitioners to address biased information in both Wikipedia and LLMs based on Wikipedia to ensure the accurate presentation of information.

Another issue is the lack of transparency regarding Wikipedia as a primary source for AI-generated content (Ford 2022; Ford and Iliadis 2023; McDowell 2024). There are growing concerns surrounding LLMs using Wikipedia as a source without proper attribution (McDowell 2024), which may potentially lead to plagiarism and copyright violations (Thomas 2023) and a disconnect between users and Wikipedia's transparent editing process. McDowell (2024) makes it clear that "LLMs often use Wikipedia as a source without acknowledging it, creating a disconnect between users and Wikipedia's rich framework" (251). As such, this lack of attribution may exert a detrimental impact on information literacy, Wikipedia's sustainability, and access to current information. This lack of transparency raises concerns regarding information literacy, as users may not be aware of the origin or reliability of the AI-generated content. For instance, many LLMs fail to cite Wikipedia, despite relying heavily on it. According to Ford and Iliadis (2023), "As Wikidata's content is ingested by knowledge graphs that power these applications, they merge data from different sources, lose the traces of their originating statements, and start to learn independently, generating new content for themselves" (9). This lack of transparency is problematic because it threatens the sustainability of Wikipedia's contributions and hinders LLM users' ability to verify information.

Finally, labor exploitation emerges as a concern regarding the use of Wikipedia as a training set. Ford and Iliadis (2023) critique the lack of consent and compensation for Wikipedia editors whose contributions are used to train commercially profitable AI models. This issue raises questions about data ownership, exploitation, and the sustainability of volunteer-driven knowledge resources. Echoing Ford and Iliadis's (2023) views, McDowell (2024) clarifies that "Wikipedia is based on explicit forms of participation in the project in comparison with LLMs extractive, nonconsensual, and often exploitative forms of inclusion into their training data" (752).

4 Feminist posthumanism and/in critical data studies

This study addresses concerns about Wikipedia as a knowledge ecosystem in the age of LLMs, particularly in regard to information biases and power dynamics embedded in the data curation processes of Wikipedia and its subsequent use in LLMs. Our theoretical approach to these concerns is informed by the principles of feminist posthumanism (Haraway 1988, 1991) and critical data studies (boyd and Crawford 2012; Crawford 2021; D'Ignazio and Klein 2020; Iliadis and Russo 2016), emphasizing the need for equitable and ethical approaches to technology's relationship to society. As such, our approach to this research is to both situate our study subjects (Wikipedia community leaders) as well as our methodological approach (problem-centered expert interview) within this lens, allowing for a critical and reflexive framework for uncovering insights into a complex and interconnected information ecosystem.

Central to the feminist posthumanist framework is to recognize the situated nature of knowledge (Haraway 1988), which posits that all knowledge is partial, contextual, and shaped by the positionality of the knower. Situated knowledge provides a methodological foundation for examining Wikipedia, not as a neutral repository of facts, but as a socially and politically embedded locus of knowledge production. As a theoretical lens, situated knowledge helps us understand and contextualize Wikipedia and its content, as well as our interviewees. We see them not only as experts and leaders in their fields, but also as Wikipedia community members, researchers, and participants in the broader information ecosystem.

It is important to remember that Wikipedia's content is heavily influenced by its predominantly male, Western contributor base, leading to significant gaps in representation for women and minorities, global perspectives, and non-Western epistemologies (Gruwell 2015; Jiang and Vetter 2020a, 2020b; McDowell 2021, 2024; Vetter et al. 2024a, 2018). By foregrounding situatedness, we can both understand and critically analyze how systemic inequalities and power asymmetries influence the creation, curation, and dissemination of knowledge on Wikipedia (and how these biases extend into AI systems trained on Wikipedia's data), as well as recognize and understand the perspectives of those close to this system and how they are keenly aware of these concerns.

Furthermore, it is important to recognize that many of the advanced algorithmic systems have historically required the "ghost work" of human labor (Gray and Suri 2019: 6) from often underpaid and underrepresented groups. Crawford (2021) echoes these concerns, noting that AI's costs disproportionately affect marginalized communities and calling for a more equitable distribution of benefits and burdens. Emerging work in critical data studies (e.g., boyd and Crawford 2012; Crawford 2021; Iliadis and Russo 2016) offer additional insights into how data archives are constructed, maintained, and leveraged, calling attention to whose knowledge is represented, whose is excluded, and how these choices reinforce existing hierarchies.

Overall, we approach this study by examining both Wikipedia and the interviewee responses within a comprehensive framework built on feminist posthumanism, critical data studies, and community (situated) knowledge. Such a framework embraces subjectivity (even bias), aiming to understand perspectives from embedded community members who recognize the complex issues threatening Wikipedia and its vital role in the information ecosystem. It also takes seriously those closest to the problem, striving to understand perspectives through and with their situations and contexts.

4.1 Methodology

This IRB-approved study (Log#: 24-072-IUP) employed a problem-centered expert interview (Döringer 2020) to investigate the relationship between Wikipedia and large language models (LLMs) as it pertains to issues of sustainability, information access and literacy, problematic information, and ethical concerns. Problem-centered expert interviews, according to Döringer (2020), involve a combination of two long-standing approaches to qualitative research, namely the broader "theory-generating expert interview" (Bogner and Menz 2009, 2018) and the "problem-centered interview" (Murray 2016; Shirani 2015; Witzel 1982, 2000). Much aligned with "situated knowledge," Döringer (2020) notes that "these epistemological perspectives… tak[e] into account their personal opinions and experiences" (269). Important to the problem-centered expert interview, for Döringer (2020), are seven (7) features and/or processes: the definition and discussion of the meaning of "expert," distinguishing "different types of expert knowledge," the goal of "inductive theory development," emphasis on "individual perspective," the use of a "specific interview design and set of questions," the capacity for comparing results, and the introduction of "inductive–deductive theory building" as shown in Table 1.

As Wikipedia is not just a knowledge repository, but a community, we employed this methodology not only to focus on "experts" in a technical dimension, but also to engage with the expertise of thought leaders within this community. These experts speak to local, embedded leadership within that community, enabling us to build a community-based epistemology. Therefore, for the purposes of our study, we use the term "expert" to not only indicate an individual with highly specialized knowledge and interest in the overlapping relationships between Wikipedia (already a very specialized subject in and of itself) and LLMs, but also as a local thought leader who brings contextual understanding grounded within that participatory community.

We also differentiate between the *types* of expert knowledge, acknowledging that both the interaction and relationships of Wikipedia with LLMs may be understood across various domains of study (computer science, natural language processing, yes, but also law, economics, and new media studies). Accordingly, expert knowledge

Table 1 Elements of the problem-centered expert interview (Döringer, 2020)

Theory-generating expert interview	Problem-centered interview (PCI)
Defines and discusses the term 'expert'	Highlights the individual perspective
Distinguishes different types of expert knowledge	Provides a specific interview design and set of questions
Aims at inductive theory development	Enables comparability of the gathered data
	Proposes inductive–deductive theory building

on the topic may be expressed as it concerns technical, social, cultural, economic, educational, or other dimensions, given the rapid acceleration of LLMs and their broad impact. The interview instrument (Appendix A) comprised eight interview questions, and the procedure itself was semi-structured to allow for follow-up questions and side discussions among interviewee and researchers. In limiting the number of participants to six, this study sought to enable comparability of the gathered data, to note when and where experts independently converged (or diverged). The methodology overall enabled both deductive and inductive theory building on topics related to the intersection of LLMs and Wikipedia, while also providing qualitative date to support and contextualize previous research (Anderl et al. 2024; Ashkinaze et al. 2024; McDowell 2024; Huang and Chang 2023; Huang and Sidarth 2023).

4.2 Recruitment

Utilizing purposive sampling, six expert participants for this study were invited based on the researchers' previous knowledge of their professional work and expertise in machine learning, Wikimedia projects, LLMs, and data science. Accordingly, all participants were both highly educated and well versed on the issues at hand, with particular insights and/or insider knowledge.

With over 25 combined years of involvement and research with Wikipedia and the Wikipedia community, the research team first identified a list of community and thought leaders within the Wikimedia movement. The research team then emailed these experts directly to solicit an interview, while also soliciting additional experts through a broader call for participants posted to the Wikimedia research listserv (wiki-research-l@lists.wikimedia.org). The research team then selected the final participants according to expertise.

As part of the informed consent process, and to best accommodate their schedules, prospective participants were given the option of a synchronous video conference interview (conducted in Zoom) or an asynchronous format conducted via email and shared cloud document (Google docs). Participants split evenly between asynchronous and synchronous formats. Participants did not receive any incentive for participating in this study.

4.3 Confidentiality

Participants were also given the option of choosing to be identified or pseudonymous as it relates to the data shared in this article. Four participants chose to be named, while two chose to remain pseudonymous, as seen below.

4.4 Expert interviewees (EIs)

1. E1 Denny Vrandečić (named)—Head of Special Projects at Wikimedia Foundation
    a. Long-term Wikimedian, computer scientist, data ontologist, and lead architect of Wikidata, the semantic database that acts as the backbone of all Wikimedia projects.
2. E2 Stephen Harrison (named)—Assistant General Counsel at *Shutterfly*
    a. Long-term Wikimedian, technology lawyer, and journalist. Significantly covered Wikipedia in *Slate*, *New York Times, Washington Post, Wired, The Guardian*, and author of *The Editors.*
3. E3 Aaron Halfaker (named)—Principal Applied Research Scientist at Microsoft
    a. Long-term Wikimedian, former research scientist at WMF, machine learning and AI researcher and developer.
4. E4 Luca de Alfaro (named)—Professor of Computer Science and Engineering, University of California Santa Cruz
    a. Long-term Wikimedian, researcher of machine learning and AI
5. E5 Ava (pseudonym)—Wikimedia researcher
    a. Long-term Wikimedian, research scientist on Wikimedia projects for over a decade.
6. E6 Andrew (pseudonym)—technical staff/researcher at WMF
    a. Long-term Wikimedian, research scientist, human–computer interaction (HCI) specialist.

4.5 Qualitative analysis

Analysis of the interview data followed a systematic six-step process and involved all members of the research group to best extract meaningful and valid conclusions via thematic analysis (Boyatzis 1998; Saldaña 2021). We used qualitative thematic analysis methods (Saldaña 2021) to code and analyze the interview data. This procedure involved assigning descriptive codes to sections of the text, organizing similar codes through axial coding, identifying core categories and themes, and refining codes and themes. After independently coding a portion of the data, we compared the results and resolved any differences through discussion, before we reached a consensus on how the codes should be applied. The following demonstrates our procedure:

1. Transcription and de-identification: for videoconference interviews, we used Zoom's automatic transcription tool for generating initial interview transcripts. Three interviews conducted via Google docs did not require transcription. Participants who requested that they remain pseudonymous were de-identified.
2. Familiarization: all members of the research group read and re-read all interview transcripts and documents.
3. Initial coding: collaborative initial coding was done in a shared Google doc.
4. Thematic analysis: we analyzed the data through open coding, axial coding, and selective coding using free and open-source software Taguette (taguette.org).
5. Coding and theme refinement: we refined the codes and themes based on our theoretical and methodological frameworks.
6. Selecting salient quotations: finally, we highlighted key quotations to report in the article, but did not attribute speakers directly in the results section.

Focusing on the experiences of the expert interviewees in this study, we coded a total of 204 units and sorted them into 16 categories. Figure 1 presents the categories from the highest to lowest frequencies of codes contained in them. The top categories include:

Sustainability challenges (37): the challenges that LLMs and LLM applications may pose to Wikipedia's long-term sustainability and maintenance.
Information literacy (30): both AI's threat to epistemology and the hope that AI will be able to use references to improve information literacy.
General ethical issues (21): general ethical questions and concerns—example: "I'm worried about humanity".

After developing first-cycle codes, we identified broader categories represented as key findings including (1) Wikipedia plays a significant role in the training of LLMs, but the exact process and value it is given is unclear; (2) LLMs act as intermediaries between users and original knowledge sources; (3) Wikipedia's sustainability is threatened by LLMs' negative impact on the digital commons; (4) The use of Wikipedia as LLM training data involves ethical problems; (5) ethical concerns may be partially addressed; and (6) systemic biases in LLMs, which can be inherited from sources like Wikipedia, are inevitable, but can be mitigated, etc. through a second round of coding. To illustrate these themes, we include the participants' quotes that reflect their perceptions of Wikipedia and LLM. Participants names were deliberately left off quoted material which were attributed to "Expert 1–Expert 6," abbreviated as "E1–E6."

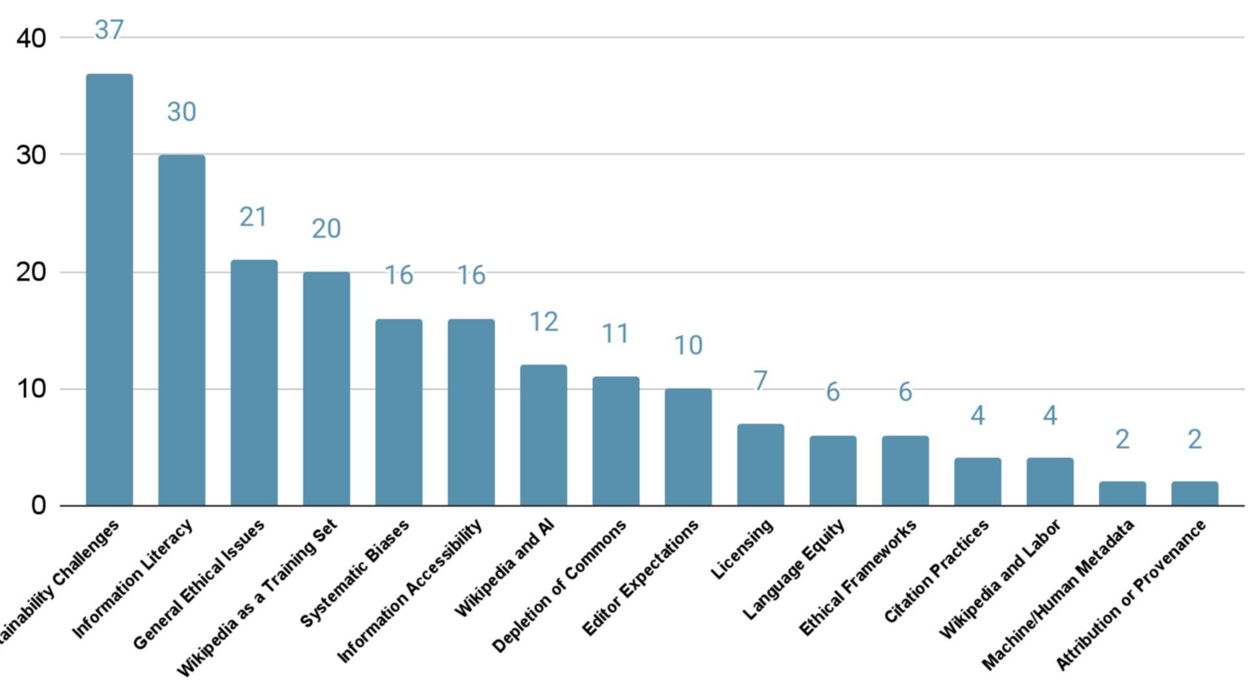

Fig. 1 Frequency count per coding category

4.6 Reliability and situated knowledge

Expert interviewees in this study were chosen due to their (1) experience with and time devoted to Wikimedia projects and (2) highly specialized knowledge related to the intersection between Wikipedia and LLMs. While two of the interviewees chose to remain pseudonymous, other participants disclosed their identities. All are highly educated with backgrounds in computer science, machine learning, and law, and all have been involved in Wikipedia for many years. Furthermore, understanding these findings through "situated knowledge" (Haraway 1988) helps to contextualize their responses, as the knowledge they derive their answers from is shaped by their unique experiences, perspectives, position within a particular field of expertise, and their membership and leadership roles within the Wikipedia knowledge community.

We found inter-expert agreement in three broad areas as part of the thematic analysis (sustainability concerns, information literacy issues, and general ethical problems) as well as more specific findings such as the use of Wikipedia in training and fine-tuning LLMs, the negative impact of LLMs on Wikipedia's discoverability, and the presence of systemic biases in LLMs, which can be partially attributed to the use of Wikipedia as a training source. Our expert participants ultimately provide knowledgeable perspectives on more specific concerns related to information ecosystems, which echo broader concerns of the community (Deckelman 2023).

4.7 Limitations

While the small sample size ($n = 6$) impacts thematic saturation, this is expected with problem-centered expert interviews, and may be a misleading, and arbitrary goal (Tight 2024). Instead, the goal in this study was to gain deeper insights of the experts' partial and situated knowledge on a future-oriented—even speculative—subject (Wikipedia's sustainability), privileging expertise within an entrenched community and local consensus over universality.

However, despite clear consensus among the participants regarding the significant role Wikipedia plays in training LLMs (see key finding 1 below), we were unable to answer one of our primary research questions regarding the specifics of how Wikipedia is actually utilized in the training of LLMs. Considering the makeup of the interviewees, including a leading expert in AI training and development, the fact that no expert was able to answer this question underscores a main issue with how AI training functions are obscured, what is sometimes called the black box problem.

5 Results

Key finding 1: Wikipedia plays a significant role in the training of LLMs, but the exact process and value it is given is unclear.

There is a clear consensus among the interviewees that Wikipedia plays a significant role in training and fine-tuning LLMs (E1–E6). For instance, the research participants noted that Wikipedia constitutes a central part of the dataset that underpins popular models such as ChatGPT and Gemini. Many expert participants emphasized that due to its open license and perceived quality, Wikipedia content is likely given more value or weight during the training process. Wikipedia may be weighted more heavily both intentionally and unintentionally (E1). According to one expert, "My understanding is that Wikipedia is intentionally given a much higher weight than many other sources. Wikipedia probably unintentionally gets an even higher weight because it's actually copied inside the web corpus several times likely." (E1). The prominence of Wikipedia as a training source, combined with its widespread availability across different online platforms, could lead to its overrepresentation in LLM training data (E1). Using a vivid metaphor, another interviewee drew attention to the way in which Wikipedia is integrated into the vast corpus of training data for LLMs: "The popular, non-technical analogy is that the training data for an LLM is like a giant hairball. Wikipedia becomes part of the hairball because it is openly licensed content." (E2). His comment implies that LLMs do not distinguish whether a piece of information originates from Wikipedia or another source, which complicates the user's ability to trace the origin of the information generated by the model.

While Wikipedia is undoubtedly a valuable resource, its predominant use in model training without clear attribution suggests the need for more transparency in how LLMs handle and prioritize various sources. Another expert highlighted how Wikipedia content is processed before being fed into LLMs: "The content of Wikipedia is surely being 'cleaned' (of some metadata) and fed into the language models that are at the basis of ChatGPT and Gemini." (E4). As a curated and structured source, Wikipedia can be optimized for language models by removing irrelevant metadata, which makes it more suitable for training. However, the lack of transparency about how data is being processed raises questions about the information being fed into LLMs (E4). Other experts also noted that although the exact process is unclear, a general procedure can be speculated: "[I]n practical terms, generally…what people are doing is they're throwing huge amounts of corpus at these models and then trying to…clean up and redirect it afterwards. And so I would suspect that they would throw the entire corpus of Wikipedia at the model. But then they might tune based on…quality assessments…. But yeah, …hard for me to say, because they don't…generally communicate about these things. But in theory this should…be likely and effective." (E3). This comment suggests that Wikipedia's content might be prioritized during later stages of model refinement, due to its quality standards, but the overall opacity surrounding LLM training practices leaves uncertainties.

Key finding 2: LLMs act as intermediaries between users and original knowledge sources, often reducing information quality and perpetuating biases, while lacking transparency and proper citation.

Although not all of our expert interviewees used the term 'dis/intermediation,' they all discussed how LLMs act as intermediaries between end users and original knowledge sources, negatively impacting both information access and information literacy (E1-E6). Referencing the function of Google's knowledge graph in Google search, one expert made a succinct point in saying that "LLM applications bring even stronger (dis-)intermediation than the Google Knowledge Panel because they are heavily customized to the question being asked." (E1). Such dis/intermediation can mean that Wikipedia is bypassed altogether due to LLMs, but also that the information quality itself suffers, whether that be in terms of simplification via a shortened summary or more problematic inaccuracy. LLMs are also prone to "hallucinations" in which they generate plausible-sounding, but inaccurate or unverified information: "It seems that the amount of misinformation coming into the system through this channel is considerably higher than it used to be." (E1). The risks of misinformation, compounded by the lack of direct access to sources, raise questions about the reliability of knowledge produced by LLMs.

On a broader level, such disintermediation also widens the already distant gap between the source of information and the original research. LLMs can provide answers to user queries, but fail to offer transparency about their sources: "LLMs often do not cite a source in their responses. Without provenance, it is difficult for the user to determine the veracity of the information." (E2). LLMs trained on Wikipedia might provide an answer to a query, but the user (often) has no access to the original source of information, the secondary source cited in Wikipedia, or even Wikipedia itself (which would already act as a distant, tertiary source). In this way, LLMs serve as quaternary sources, three times removed from the original production of the information or knowledge. This distance is further explained as a gap between source and consumption: "The distance between the source, both in the cases of computing technology as well as original research, and the consumption of it is…a concerning gap." (E3). LLMs, especially those trained on publicly available, tertiary content like Wikipedia, can both negatively impact the accuracy of information and further disintermediate users from the original knowledge creation process.

Key finding 3: Wikipedia's sustainability is threatened by LLMs' negative impact on the digital commons, Wikipedia discoverability, community engagement, and disintermediation.

If LLMs are acting as intermediaries and directing traffic away from the actual encyclopedia (while relying on training data from the encyclopedia), how might their development affect Wikipedia's long-term sustainability? The interviewees expressed concerns regarding LLMs having negative impact on the digital commons, discoverability, community participation and engagement (attracting new editors), and disintermediation. The risk of a shrinking open (commons) environment especially could isolate Wikipedia and hinder its collaborative nature (E5). Users may rely on LLMs for quick consultations, bypassing Wikipedia and reducing opportunities for content improvement and community engagement (E5). All of the interviewees warn that LLMs could diminish the discoverability of Wikipedia, leading to decreased donations and editorial contributions (E1–E6). One expert recommended Wikipedia should position itself as a crucial resource for training LLMs as a way to attract new contributors, but also noted the risk of LLMs overshadowing human-generated content (E2). Additional emphasis was placed on the importance of maintaining Wikipedia's feedback loop, where readers become contributors, and caution against tools that replace rather than support Wikipedians (E3). Because disintermediation could undermine the motivation for community engagement, there is a need for targeted outreach via WikiProjects and campaigns in fostering a diverse and engaged editor community (E6). The same expert also stressed the necessity of making sources easier to work with to ensure high-quality content and suggested integrating AI-supported content with traditional human-written content to enhance accessibility (E6). Ultimately, the sustainability of Wikipedia depends on continuous experimentation and technical support to adapt to the evolving digital landscape as it is disrupted by emerging generative AI and LLM tools (E1, E6).

A related danger, though only expressed by one participant, is the potential for a competitor to emerge, using LLMs to create personalized content, thereby drawing users away from Wikipedia and undermining its foundational community. Wikipedia's unique, non-profit model is crucial for its survival, as it deters commercial competitors from attempting to replace it (E1). Once lost, Wikipedia's collaborative and comprehensive knowledge base would be nearly impossible to recreate, given the historical and communal efforts that built it (E1). This underscores the importance of maintaining Wikipedia's role as a primary knowledge source to prevent the erosion of its community and the valuable content it provides.

Key finding 4: the use of Wikipedia as LLM training data involves ethical problems related to contributor expectations, the risk of depleting the commons, and exacerbation of linguistic and cultural inequities.

Interview participants were asked to respond to the following questions regarding ethical concerns: "In your opinion, what ethical problems or issues, if any, emerge in terms of the relationship between Wikipedia and its use as training data for LLMs?" All but one interviewee agreed that this relationship constituted an ethical problem, and responses were categorized in the following themes: contributor expectations, risks to the digital commons, and linguistic and cultural inequities.

There is agreement among expert interviewees that Wikipedia contributors never intended for their content to be used by machine learning models (E2, E4). "The fundamental problem, as one expert puts it, "is that users that would have been quite happy to provide their content to other humans, are not necessarily happy to have their content fed to a [machine learning] model. That is, when determining licensing rights, it seems that the current body of law makes the glaring omission of not mentioning, in the license, the expected and intended audience, at the time, for the licensing." (E4). Another interviewee echoes this sentiment, noting that many Wikipedians feel it is unfair that their unpaid work is used by big tech companies to generate profit: "The ethical problem that I hear about most frequently from Wikipedians is that the situation doesn't seem fundamentally fair. The editors produce this content without compensation, it is openly licensed, and then these big tech companies make so much money from LLMs." (E2).

Another central concern among our EIs is that the overuse of digital commons content for training LLMs could deplete the commons by exhausting available resources and discouraging contributors who feel their work is exploited without recognition or compensation (E2, E5). The current AI race, with multiple tech companies competing to develop and fine-tune LLMs, further exacerbates this issue, as does the fact that there has been no attention to reciprocity (or giving back to) the commons (E5). There is an ethical obligation to give back to the commons proportionately to what is extracted, stressing the importance of maintaining the sustainability of these shared resources (E5). Other participants concur with the need for giving back, suggesting that human-generated content will become increasingly valuable as it becomes rarer (E2).

Expert interviewees also expressed significant concerns regarding the ethical implications of LLMs on linguistic and cultural (in)equities, especially when it comes to access and representation (E1, E3, E5, E6). Because Wikipedia already relies on and extends English as a dominant language, training LLMs on this data highlights the risk of exacerbating existing gaps in access to technology and the Internet, particularly for speakers of less dominant languages. LLMs are limited in multiple languages due to the high costs of running these models, which raises questions about scalability and inclusivity (E5). LLMs, like Wikipedia, rely heavily on digitized documents, which exist mostly in dominant languages (E3). This reliance can marginalize cultures with less digital documentation, potentially leading to cultural erasure (E3). As one expert states, "[T]here's a concern around equity—leaving people behind or forcing people to [use] languages that [are not their] native languages. They are the languages of the colonizers." (E5). To make matters worse, LLMs perform well with widely documented languages but struggle with less common ones, further entrenching systemic biases (E3). Ultimately, the language modeling community urgently needs to address these challenges to prevent long-term consequences and ensure broader language coverage and representation (E6).

Key finding 5: ethical concerns may be partially addressed via systemic changes to market incentives and license models, financial contributions to Wikipedia from big tech, and technical solutions related to data provenance and attribution.

While the existence of ethical issues as it relates to Wikipedia being used as a training data was not agreed upon unanimously, a majority of experts both identified ethical issues and proposed possible solutions to address such issues, proposing a variety of fixes related to licensing, market incentives, LLM explainability, and data provenance.

On a broader scale, there is a need for a radical rethinking of market incentives and licensing models to ensure the sustainability of the digital commons (E5). One expert references Larry Lessig's work on redesigning market incentives (Lessig 2022), arguing that profit maximization should not be the sole reward mechanism (E5). Wikipedia has long thrived on the altruism of its volunteer contributors, but that model is endangered by the LLM economy in which digital commons content is extracted and exploited beyond the expectations of its original creators, and without respecting CC-BY-SA licensing. In contrast to this emphasis on market incentives, another interviewee calls for immediate financial contributions from big tech as a necessary step to support the commons (E2). "Big tech should contribute to the project," this expert notes, "but it is very important that big tech does not itself have any editorial influence." (E2).

The role of Wikipedia in this context is also a point of contention. While Wikipedia can contribute to the broader open-source movement, it is not solely responsible for solving the open culture challenge (E5). An online encyclopedia's primary role is not to address these issues of open source and open culture, although it can play a supportive role (E5). One way that LLMs might address issues related to information literacy loss among users, for example, is the addition of explainability measures, which was frequently referenced by one expert. Such explainability would ensure that LLMs describe to users how and where they retrieved certain information or outputs (E3). Noting the opportunities in training LLMs to express "chain of thought", this expert expressed how LLMs might showcase "processes that would probably look very familiar to Wikipedia and information literacy processes." (E3). If developers focused less on the speed of outputs and emphasized "quality and information literacy instead," we might end up with a model that is "able to talk to you about what it's doing and what it's thinking." (E3).

Finally, technical solutions related to data provenance and attribution, such as ensuring LLMs include citations, are necessary to maintain the integrity of the commons (E2). Going forward, human-generated content will be considered even more valuable as it becomes increasingly rare (E2). While there is a shared concern about the depletion of the commons and the need for giving back, the participants differ in their approaches to addressing these issues, with some experts advocating for systemic changes to market incentives and licensing models, while others emphasize immediate financial contributions and technical solutions to maintain the integrity of the commons.

Key finding 6: systemic biases in LLMs, which can be inherited from sources like Wikipedia, are inevitable, but can be mitigated via proactive efforts to diversify communities and content in the digital commons.

EIs collectively highlighted the pervasive issue of systemic biases in LLMs and their potential perpetuation from sources like Wikipedia (E1–E6). As one expert stated, "Yes, there is a risk of these systemic biases being perpetuated in LLMs. To the extent there are systemic biases in Wikipedia, or the broader media landscape, then it is likely that the LLMs will be trained on these same biases" (E6).

While the encyclopedia itself has improved (and can continue to improve) by proactive efforts by Wikipedia editors to address bias through dedicated task forces, there are inherent biases due to limited content in various languages (E5). This is inevitable, reflecting the human biases of contributors, and actively including more diverse communities can mitigate these effects (E5). Additional EIs affirm the risk of systemic biases in LLMs (E2, E4), with one pointing out that these biases are likely to be inherited from the media (E2). Another expert discussed the dominance of Western documentation practices in Wikipedia, which can marginalize non-Western knowledge systems, and underscored the need for diverse sources to avoid cultural erasure. As noted by this expert, "Western culture has really strongly adopted this whole documentary practice around knowledge, and that fits with Wikipedia. But there's all sorts of knowledge all around the world that aren't documented in familiar ways, or maybe aren't documented." (E3). The issue of language diversity further compounds the issue: "A lack of language coverage (and therefore perspectives from these other language communities) is probably the most concerning aspect of bias to me with these models." (E6). Despite the fact that Wikipedia does better than much of the Internet in offering multilingual content, significant linguistic gaps still exist on Wikipedia, especially in underrepresented languages and communities. This lack of linguistic diversity in Wikipedia is mirrored in LLMs, which are disproportionately trained on dominant languages with a lack of representation of non-Western knowledge systems (E6). Finally, one of the most obvious examples of systemic biases in LLMs are issues in translation systems (E1). Overall, biases in training data are almost certain to appear in LLMs unless explicit efforts are made to counteract them (E1).

6 Implications of results

Based on the findings from this study, we identified three main areas of opportunity. First, advocating for transparency and attribution in LLM applications, especially those that benefit from training on Wikipedia, would benefit the health of the entire information ecosystem. Doing so will direct traffic back to the encyclopedia, while also demonstrating Wikipedia's importance when it comes to information reliability. Attribution of Wikipedia, furthermore, acknowledges the contributions of the community and maintains the integrity of information. Along the same lines, encouraging the development of LLMs that prioritize explainability and transparency, could provide users with insights into how they retrieve and process information. For example, data scientists, researchers, and machine learning specialists could advocate for explainable AI (XAI), which aims to make AI systems more understandable to humans and which is essential for addressing concerns related to bias, misinformation, accountability, and data provenance (Kale et al. 2023).

Second, while attribution will help to redirect traffic and give credit back to Wikipedia contributors, there remains a need to continue to improve the encyclopedia as it feeds LLMs the most accurate and comprehensive information. Supporting initiatives to diversify the Wikipedia editor community and recruit new users and contributors helps to encourage a plethora of contributions, particularly from underrepresented groups and languages. The Wikimedia community is already doing this, of course ("Movement Strategy/Recommendations/Increase" 2024). But these initiatives take on heightened importance given the way that Wikipedia content is utilized in downstream applications such as LLMs. This type of work could also include campaigns for raising awareness about the value of Wikimedia projects—for example, promoting Wikipedia's unique attributes such as its collaborative editing, neutrality, and trusted information can help recruit users and contributors and ensure its continued relevance and discoverability in the age of LLMs.

Third and finally, the concerns related to information literacy among our experts necessitate new and continued educational initiatives that equip users with the critical thinking skills necessary to evaluate AI-generated content and understand its limitations, as well as to understand Wikipedia's role in the evolving information landscape. This means designing new pedagogical interventions aimed at helping users understand and negotiate the complex intersections between AI-generated content produced via LLMs and other information sources, especially Wikipedia and other Wikimedia projects. Allowing for and funding professional development for educators at the postsecondary level (and really all levels) will help to build the necessary skills to engage students in critical thinking activities related to LLMs and information literacy.

7 Discussion

While our key findings provide significant takeaways, this research is complicated by the ongoing and unresolved nature of the subject. Our interview questions required EIs to speculate on numerous unknowns, which, while illuminating as a wayfinder due to their expertise and experience, remains tenuous in the ever-changing landscape of AI. This being said, what remains unknown to even experts and developers helps to frame the significance of the unknown, particularly regarding training "black boxes." What has emerged here reinforces previous articulations and concerns as well as provides some challenges for future research on AI, both from a development and an ethical perspective.

First of all, the findings of this study resonate with the posthumanist approach to critical AI literacy (Burriss and Leander 2024; Jiang 2024; Leander and Burriss 2020; Vetter et al. 2024a, b). Going beyond simply identifying computational and algorithmic agents, the critical posthumanist approach emphasizes the active construction of entanglements with these nonhuman agents (Jiang et al. 2024; Leander and Burriss 2020; Vetter et al. 2024a, b). Leander and Burriss (2024) use the term "sociotechnical justice" (7) to foreground the entanglement of humans and machines in the process of producing equity in social systems. The entangled view of power in sociotechnical justice aligns with the concept of AI as a hyperobject (Burriss and Leander 2024), which suggests that the development of ethical practices is a dynamic process that is constantly evolving in response to humans' interactions with technological agents. Our understanding of AI ethics contributes to the ongoing negotiation of power dynamics when using Wikipedia as an AI training set, where the lines between human and machine contributions are becoming increasingly blurry. The EI interviews in this study illustrate the development of AI ethics by at once showing how LLMs may pose a threat to Wikipedia's sustainability and demonstrating how proactive efforts can mitigate and address this ethical concern.

While research is ongoing and there is a general sense of positivity towards the potential applications and uses for LLM systems (especially utilizing new systems to fact check or link back to Wikipedia, or new requirements for LLMs to provide sources), overall there remains a deep concern among these experts for (continued) misuse and erosion of both information literacy and also for the future of Wikipedia and the digital commons. While expert opinions mirror some concerns raised by previous research (Anderl et al. 2024; Ford 2022; McDowell 2024; McDowell and Vetter 2022; Reeves et al. 2024), particularly regarding community engagement, systemic biases, information literacy, sourcing, and exacerbating existing accessibility concerns, they also provided some potentials for alleviation. In particular, the researchers who worked on these models noted that some of these concerns (sourcing in particular) were driving new advancements that would shape the front-end experience with LLMs to include things like sourcing so that some level of context might be included in LLM responses. As models continue to be further developed, Wikimedia stakeholders and the digital commons should continue to advocate for explainability as a necessary component for LLMs and generative AI in the context of information literacy.

Overall, the EIs tended to retain hope that the technical concerns could be addressed through technical means, while also attending to communities of users. While the particularities of *how* might have changed, the hope for technological fixes seem to persist despite the growing cynicism. Whether this approach will bear fruit, like many of the pending questions regarding LLMs, is yet to be seen.

Of course, the marketplace remains the greatest concern of all. Wikipedia's success relies on the donation of labor by its volunteers, which runs counter to the rest of the Internet. Market incentives and profit maximization continue to threaten the sustainability of the digital commons as these systems continue to overtake other repositories of knowledge and representation. This tenuous relationship may force the market to rethink its approach, or find an alternate relationship. However, in the end, big tech companies rely on Wikipedia, and, at least for now, seem to have a vested interest in its continuation. As one interview put it, "No one wants to come in and kill Wikipedia" and be known for that (E1). If nothing else, this marriage of convenience, when combined with Wikipedia's long-standing nonprofit and altruistic status, may spark a glimmer of hope for those concerned with the future.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00146-025-02199-9.

Acknowledgements The authors would like to thank interview participants for contributing their time and expertise to this study.

Author contributions All authors conceptualized the research design, analyzed data, and contributed to the drafting of the the article. MV arranged and conducted IRB approval and interviews with participants.

Data availability No datasets were generated or analysed during the current study.

Declarations

Conflict of interest The authors declare no competing interests.

References

Anderl C, Klein SH, Sarigül B, Schneider FM, Han J, Fiedler PL, Utz S (2024) Conversational presentation mode increases credibility judgements during information search with ChatGPT. Sci Rep 14(1):17127. https://doi.org/10.1038/s41598-024-67829-6

Ashkinaze J, Guan R, Kurek L, Adar E, Budak C, Gilbert E (2024) Seeing like an AI: how LLMs apply (and misapply) Wikipedia neutrality norms. arXiv. https://doi.org/10.48550/arXiv.2407.04183

Bogner A, Menz W (2009) The theory-generating expert interview: Epistemological interest, forms of knowledge, interaction. In: Interviewing experts. Palgrave Macmillan UK, London, pp 43–80

Bogner A, Menz W (2018) Collecting data with experts and elites. In: Flick U (ed) The Sage handbook of qualitative data collection. SAGE 652–667

Borgman CL, Van de Sompel H, Scharnhorst A, van den Berg H, Treloar A (2015) Who uses the digital data archive? an exploratory study of DANS. In: Proceedings of the Association for Information Science and Technology 52(1):1–4. https://doi.org/10.1002/pra2.2015.145052010096

Bowker GC (2010) The archive. Commun Crit/ Cult Stud 7(2):212–214. https://doi.org/10.1080/14791421003775733

Boyatzis RE (1998) Transforming qualitative information: thematic analysis and code development. Sage, London

Boyd D, Crawford K (2012) Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon. Inf Commun Soc 15:662–679. https://doi.org/10.1080/1369118X.2012.678878

Burriss SK, Leander K (2024) Critical posthumanist literacy: building theory for reading, writing, and living ethically with everyday artificial intelligence. Read Res Q 565:1–10. https://doi.org/10.1002/rrq.565

Buneman P, Cheney J, Lindley S, Müller H (2011) DBWiki: a structured wiki for curated data andcollaborative data management. In: Proceedings of the 2011 ACM SIGMOD International Conference onManagement of data. https://doi.org/10.1145/1989323.1989491

Cooban G (2017) Should archivists edit Wikipedia, and if so how? Arch Rec 38(2):257–272. https://doi.org/10.1080/23257962.2017.1338561

Crawford K (2021) The atlas of AI: power, politics, and the planetary costs of artificial intelligence. Yale, New Haven

D'Ignazio C, Klein LF (2020) Data feminism. MIT, Cambridge

Deckelmann S (2023) Wikipedia's value in the age of generative AI. Wikimedia Foundation. https://wikimediafoundation.org/news/2023/07/12/wikipedias-value-in-the-age-of-generative-ai/. Accessed 11 Jul 2024

Döringer S (2020) The problem-centred expert interview': combining qualitative interviewing approaches for investigating implicit expert knowledge. Int J Soc Res Methodol 24:265–278. https://doi.org/10.1080/13645579.2020.1766777

Fichman P, Hara N (eds) (2014) Global Wikipedia: international and cross-cultural issues in online collaboration. Rowman & Littlefield, Lanham, MD

Ford H (2022) Writing the revolution: Wikipedia and the survival of facts in the digital age. MIT, Cambridge

Ford H, Iliadis A (2023) Wikidata as semantic infrastructure: knowledge representation, data labor, and truth in a more-than-technical project. Soc Media Soc. https://doi.org/10.1177/20563051231195552

Fuster Morell M (2014) Governance of online creation communities for the building of digital commons: viewed through the framework of the institutional analysis and development. SSRN Electron J. https://doi.org/10.2139/ssrn.2842586

Gertner J (2023) Wikipedia's moment of truth. The New York Times. https://www.nytimes.com/2023/07/18/magazine/wikipedia-ai-chatgpt.html

Giovannoni E, Fabietti G (2013) What is sustainability? a review of the concept and its applications. In: Busco C, Frigo ML, Riccaboni A, Quattrone P (eds) Integrated reporting. Springer International Publishing, Cham, pp 21–40

Gray ML, Suri S (2019) Ghost work: how to stop silicon valley from building a new global underclass, Illustrated. Harper Business, Boston

Gruwell L (2015) Wikipedia's politics of exclusion: gender, epistemology, and feminist rhetorical (in) action. Comput Comp 37:117–131. https://doi.org/10.1016/j.compcom.2015.06.009

Haraway D (1988) Situated knowledges: the science question in feminism and the privilege of partial perspective. Fem Stud 14(3):575–599. https://doi.org/10.2307/3178066

Haraway D (1991) Simians, cyborgs, and women: the reinvention of nature. Routledge, New York

Huang J, Chang KCC (2023) Citation: a key to building responsible and accountable large language models. arXiv. https://doi.org/10.48550/arXiv.2307.02185

Huang S, Siddarth D (2023) Generative AI and the digital commons. arXiv. https://doi.org/10.48550/ARXIV.2303.11074

Iliadis A, Russo F (2016) Critical data studies: an introduction. Big Data Soc 3:205395171667423. https://doi.org/10.1177/2053951716674238

Jiang J, Vetter MA (2020a) Addressing the challenges and opportunities of a feminist rhetorical approach for Wikipedia-based writing instruction in first-year composition. Composition Forum 44. https://www.compositionforum.com/issue/45/wikipedia.php

Jiang J, Vetter MA (2020b) The good, the bot, and the ugly: problematic information and critical media literacy in the postdigital era. Postdigit Sci & Educ 2(1):78–94. https://doi.org/10.1007/s42438-019-00069-4

Jiang J (2024) When generative artificial intelligence meets multimodal composition: rethinking the composition process through an AI-assisted design project. Comput & Compos. https://doi.org/10.1016/j.compcom.2024.102883

Jiang J, Vetter MA, Lucia B (2024) Toward a “more-than-digital” AI literacy: reimagining agency and authorship in the postdigital era with ChatGPT. Postdigit Sci & Educ 6:922–939. https://doi.org/10.1007/s42438-024-00477-1

Kale A, Nguyen T, Harris FC, Li C, Zhang J, Ma X (2023) Provenance documentation to enable explainable and trustworthy AI: a literature review. J Data Intell 5(1):139–162. https://doi.org/10.1162/dint_a_00119

Leander KM, Burriss SK (2020) Critical literacy for a posthuman world: when people read, and become, with machines. Br J of Educ Technol 51(4):1262–1276. https://doi.org/10.1111/bjet.12924

Lessig L (2022) Keynote: How can the Internet be so bad and so good: the lessons we must draw and that Wiki must teach. Wiki Workshop 2022. https://www.youtube.com/watch?v=LlWOt7jqT1M. Accessed 17 Sep 2024.

Lichtenstein S, Parker CM (2009) Wikipedia model for collective intelligence: a review of information quality. Int J Knowl Learn 5(3/4):254. https://doi.org/10.1504/IJKL.2009.031199

Liu Y, Cao J, Liu C, Ding K, Jin L (2024) Datasets for large language models: a comprehensive survey. arXiv. https://doi.org/10.48550/ARXIV.2402.18041

McDowell ZJV, MA, (2021) Wikipedia and the representation of reality. Routledge, New York

McDowell ZJ, Vetter MA (2022) Fast “truths” and slow knowledge; oracular answers and Wikipedia’s epistemology. Fast Capital 19(1):104–112. https://doi.org/10.32855/fcapital.202201.009

McDowell ZJ (2024) Wikipedia and AI: access, representation, and advocacy in the age of large language models. Convergence 30:751–767. https://doi.org/10.1177/13548565241238924

McDowell, ZJ Vetter MA (2024) The Re-alienation of the commons: Wikidata and the ethics of “free” data. Int J Commun 18:590–608. https://ijoc.org/index.php/ijoc/article/view/20807

Movement strategy/Recommendations/Increase the sustainability of our movement - Meta. Wikimedia Foundation. 2024. https://meta.wikimedia.org/wiki/Movement_Strategy/Recommendations/Increase_the_Sustainability_of_Our_Movement. Accessed 16 Dec 2024

Murray LC (2016) Book review: the problem-centred interview. J Mix Methods Res 10(1):112-113.https://doi.org/10.1177/1558689815577032

Pestoff V (2014) Collective action and the sustainability of co-production. Public Manag Rev 16(3):383–401. https://doi.org/10.1080/14719037.2013.841460

Petroni F, Broscheit S, Piktus A, et al. (2023) Improving Wikipedia verifiability with AI. Nat Mach Intell 5:1142–1148. https://doi.org/10.1038/s42256-023-00726-1

Redi M, Fetahu B, Morgan J, Taraborelli D (2019) Citation needed: a taxonomy and algorithmic assessment of Wikipedia’s verifiability. In: The World Wide Web Conference. pp. 1567–1578. https://doi.org/10.1145/3308558.3313618

Reeves N, Yin W, Simperl E (2024) Exploring the impact of ChatGPT on Wikipedia engagement. arXiv. https://doi.org/10.48550/arXiv.2405.10205

Saldaña J (2021) The coding manual for qualitative researchers, 4th edn. Sage, London

Schaul K, Chen SY, Tiku N (2023) Inside the secret list of websites that make AI like ChatGPT sound smart. Washington Post. https://www.washingtonpost.com/technology/interactive/2023/ai-chatbot-learning/. Accessed 15 Sept 2024

Smith CE, Yu B, Srivastava A, Halfaker A, Terveen L, Zhu H (2020) Keeping community in the loop: understanding Wikipedia stakeholder values for machine learning-based systems. arXiv. https://doi.org/10.48550/arXiv.2001.04879

Shirani F (2015) Book review: Andreas Witzel and Herwig Reiter, the problem-centred interview. QualRes 15(1): 134–135. https://doi.org/10.1177/1468794114535037

Srinivasan K, Raman K, Chen J, Bendersky M, Najork M (2021) Wit: Wikipedia-based image text dataset for multimodal multilingual machine learning. arXiv. https://doi.org/10.48550/arXiv.2103.01913

Sugimoto S, Nagamori M, Mihara T, Honma T (2015) Metadata in cultural contexts—from manga to digital archives in linked open data environment. In: Ruthven I, Chowdhury GG (eds) Cultural heritage information: Access and management. Routledge, New York

Szajewski M. (2013) Using Wikipedia to enhance the visibility of digitized archival assets. D-Lib Mag 19(3). https://www.dlib.org/dlib/march13/szajewski/03szajewski.html

Thomas PA (2023). Wikipedia and large language models: perfect pairing or perfect storm? Libr Hi Tech News 40(10):6–8. https://kuscholarworks.ku.edu/handle/1808/34102

Tight M (2024) Saturation: an overworked and misunderstood concept? Qual Inq 30(7):577–583. https://doi.org/10.1177/10778004231183948

Vetter MA, Jiang J, Othman M, Muguimi M (2024a) Navigating the emotional terrain of Wikipedia writing: a feminist affective analysis of student writers’ engagement with the “be bold” guideline. Comput & Compos. https://doi.org/10.1016/j.compcom.2024.102850

Vetter MA, Lucia B, Jiang J, Othman M (2024b) Towards a framework for local Interrogation of AI ethics: a case study on text generators, academic integrity, and composing with ChatGPT. Comput & Compos. https://doi.org/10.1016/j.compcom.2024.102831

Vetter MA, Andelfinger J, Asadolahi S, Cui W, Jiang J, Jones T, Siddique ZF, Tanasale IO, Ylonfoun AE, Xing J (2018) Wikipedia’s gender gap and disciplinary praxis: representing women and minority scholars in digital rhetoric and writing fields. J Multimodal Rhetor 2(2). http://journalofmultimodalrhetorics.com/2-2-vetter-et-al

Wagner C, Jiang L (2025) Death by AI : will large language models diminish Wikipedia? Assoc Inf Sci & Tech. https://doi.org/10.1002/asi.24975

Wikimedia Foundation (2024). https://wikimediafoundation.org/. Accessed 17 December 2024

Wikipedia:Verifiability. Wikipedia (2024). https://en.wikipedia.org/w/index.php?title=Wikipedia:Verifiability&oldid=1244846659. Accessed 12 Sept 2024

Witzel A (1982) Verfahren der qualitativen sozialforschung: uberblick und alternativen. Campus-Verlag,Frankfurt.

Witzel A (2000) Das problemzentrierte interview. Forum Qualitative Sozialforschung/Forum: QualitativeSocial Research 1(1). https://doi.org/10.17169/fqs-1.1.1132

Wong K, Redi M, Saez-Trumper D (2021, July) Wiki-reliability: a large scale dataset for content reliability on wikipedia. arXiv:2105.04117v2. https://arxiv.org/abs/2105.04117

Zhang CC, Houtti M, Smith CE, et al (2022) Working for the invisible machines or pumping information into an empty void? An exploration of Wikidata contributors’ motivations. Proc ACM Hum Comput Interact 6:1–21. https://doi.org/10.1145/3512982