

Exploring the Spectrum between Human Editing and Bot Editing with Wikipedia Metadata

Max Wang
Clarkson University

Ashik Ahamed
Clarkson University

Jeanna Matthews
Clarkson University

Abstract

Wikipedia editing increasingly blends human judgment with automated and semi-automated tools, challenging simple binary human-versus-bot distinctions. We examine whether Wikipedia’s Special:Tags metadata can better capture the complexity of the spectrum between human editing and bot editing. By analyzing tag definitions and revision-level data, we show that many edits fall between fully manual and fully automated actions. Our findings suggest that metadata provide a transparent way to study hybrid authorship and the evolving role of automation in collaborative knowledge production.

Introduction

Wikipedia is often described as a collaboratively authored encyclopedia maintained by human volunteers and automated bots. Early research and community governance frameworks have commonly treated these two contributors as binary categories: edits are either made by humans or by bots (Geiger and Halfaker, 2013). However, contemporary Wikipedia editing no longer fits neatly into this simple framing. A growing number of edits are produced through semi-automated tools, assisted editing interfaces, and scripted workflows that combine human decision-making with machine execution.

This shift has been accelerated by automation tools such as AutoWikiBrowser and task-specific scripts (such as for combatting vandalism) that allow editors to perform large-scale or repetitive changes efficiently while still retaining human oversight. Generative AI tools represent an emerging frontier but remain largely outside current tagging infrastructure. At the same time, fully autonomous bots continue to perform maintenance and policy-driven tasks at scale (Niederer and van Dijck, 2010). As a result, many edits occupy an intermediate space between purely human and purely automated actions. When it comes to identifying AI-generated content, recent research has focused heavily on using linguistic features or model-based classifiers (Ippolito et al., 2020). While such approaches are valuable, they face challenges related to accuracy and consistency. In contrast, Wikipedia already records extensive metadata

about how edits are made, including tool usage and automation indicators. Consider the Special:Tags system, which attaches labels to revisions based on editing tools, interfaces, or automated processes involved (Wikipedia, 2025; Ahamed et al., 2026). In this work, we ask whether Wikipedia’s existing metadata—specifically Special:Tags—can be used to reveal a nuanced spectrum of human-to-bot involvement in text editing. Rather than proposing new methods to detect involvement of automated systems, we examine what signals are already available and how they can be interpreted to better understand hybrid authorship in collaborative knowledge production. While we try to leverage Wikipedia’s tagging infrastructure, we also acknowledge its limitations: even though it can track standard automation, it can’t necessarily distinguish generative AI workflows. Regardless, our aim is to introduce nuance into the discussion of human-versus-bot editing.

Methods

Our analysis combines qualitative examination of tag definitions with quantitative analysis of revision-level metadata. We begin by collecting the full set of tags listed on Wikipedia’s Special:Tags page, focusing on tags related to editing actions rather than purely administrative events. Each tag is reviewed based on its description, intended use, and associated tooling. We categorize tags along two primary dimensions. First, we identify the type of agent implied by the tag: human, bot, or both. Second, we assess the degree of automation involved, ranging from fully manual editing to semi-automated assistance and fully automated execution. This process results in a taxonomy that reflects a continuum of editorial agency rather than discrete categories. To validate these categorizations, we analyze revision metadata using SQL queries on Wikipedia’s public databases (Wikimedia Foundation, 2024). To audit our categorizations, we examine the account type of each tag’s earliest recorded usage and examine the associated user account to determine whether the edit was performed by a registered bot account or a human editor. We acknowledge that account type is itself a binary human/bot label—the same simplification we aim to move beyond—and therefore treat it as a consistency check rather than a measure of automation level (Geiger and Ribes, 2011). We also analyze patterns of tag usage over time and examine cases where multiple tags are applied to a single

revision, indicating layered workflows. This mixed-methods approach allows us to compare the intended meaning of tags with their actual use in practice and to clarify ambiguities in how automation is represented in Wikipedia’s metadata.

Results

Our analysis shows that Wikipedia’s Special:Tags system already encodes substantial information about the diversity of editing practices, though this information is often underutilized. Some tags clearly indicate fully automated actions performed by bot accounts, such as routine maintenance or policy enforcement tasks. At the opposite end of the spectrum, some tags correspond to fully manual edits made through standard editing interfaces. However, a large portion of commonly used tags fall between these extremes.

Tags associated with semi-automated tools often involve significant machine assistance, such as automated text transformations or batch operations, while still requiring human initiation and review. These edits cannot be accurately described as either purely human or purely bot-generated. We also find that tag usage does not always align cleanly with assumptions about the level of editorial agency based on tag descriptions. In several cases, tags categorized as bot-only are first used by human editors operating tools manually, while some tags labeled as hybrid are disproportionately associated with high-volume, automation-heavy editing. Additionally, the co-occurrence of multiple tags on a single revision highlights the complexity of modern editing workflows, where human editors interact with multiple layers of automation in producing a single change. These findings suggest that Special:Tags reflect a spectrum of human–machine collaboration rather than a binary distinction. At the same time, inconsistencies in tag naming, documentation, and usage limit their interpretability without careful analysis.

Discussion and Social Implications

The results highlight the potential of metadata-based approaches for studying authorship and automation on Wikipedia. Unlike text-based AI detection methods, metadata signals are transparent, auditable, and embedded directly in the editorial process. Interpreted carefully, they provide valuable insights into how humans and machines collaborate to produce encyclopedic content. This has important social and ethical implications. Wikipedia content is widely reused for research, education, and as training data for machine learning systems. A more nuanced understanding of editorial agency can support better dataset documentation, improved transparency, and more responsible downstream use.

From a community perspective, clearer representations of automation may also help maintain trust. Readers and editors alike benefit from understanding how content is created, especially as automation tools become more prevalent and generative AI editing remains largely untracked. However, our findings also underscore the limitations of the current tagging system. Inconsistent application and vague definitions reduce its effectiveness as a signaling mechanism, particularly for emerging AI-assisted tools that may not be explicitly tagged.

Limitations

The Special:Tags system was not designed as a research instrument, and our analysis reflects its imperfections. Tags are inconsistently applied, variably documented, and sometimes used in ways that diverge from their stated definitions. We do not treat this taxonomy as a precise classifier of editorial agency. Rather, we interpret tags as directional signals — imperfect but transparent indicators that, taken together, reveal patterns of automation that would otherwise be invisible. Future work should focus on improving the reliability of this infrastructure before it can be used for high-stakes attribution purposes.

Future Work and Conclusion

Future work should explore ways to refine and extend Wikipedia’s tagging infrastructure to better capture degrees of automation and AI involvement. This may include introducing new tags specifically designed for AI-assisted editing, which current infrastructure does not yet capture, and improving documentation for existing tags. Combining metadata infrastructure with finer-grained approaches, such as sentence-level revision tracking, could further enhance attribution and accountability. In conclusion, Wikipedia’s Special:Tags metadata offers a promising foundation for moving beyond simplistic human-versus-bot narratives. By interpreting tags as indicators along a continuum of editorial agency, we gain a more realistic understanding of hybrid authorship in collaborative knowledge production. As AI-assisted editing continues to evolve, leveraging existing metadata provides a transparent and scalable path toward greater clarity and accountability.

References

- R. Stuart Geiger and Aaron Halfaker. 2013. Using edit sessions to measure participation in Wikipedia. In Proceedings of the 2013 Conference on Computer Supported Cooperative Work, 861–870.
- R. Stuart Geiger and David Ribes. 2011. Trace ethnography: following coordination through documentary practices. In 2011 44th Hawaii International Conference on System Sciences, 1–10. IEEE.
- 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):1368–1387.
- Ashik Ahamed, Max Wang, Amity Ramona Mentis-Cort, and Jeanna Matthews. 2026. An analysis of Wikipedia's special tags and their implications for the nuanced spectrum between human edits and bot edits. *International Conference on Virtual Learning*, 21:15–26. <https://doi.org/10.58503/icvl-v21y202601>
- Daphne Ippolito, Daniel Duckworth, Chris Callison-Burch, and Douglas Eck. 2020. Automatic detection of generated text is easiest when humans are fooled. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 1808–1822.
- Wikipedia. 2025. Help:Tags. Wikimedia Foundation.
- Wikimedia Foundation. 2024. Wikimedia databases: schema and data access documentation. https://www.mediawiki.org/wiki/Manual:Database_layout

