Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

The AI Democratization Paradox: Evidence from Decentralized Knowledge Communities

Kai Zhu, Dylan Walker

To cite this article:
Kai Zhu, Dylan Walker (2026) The AI Democratization Paradox: Evidence from Decentralized Knowledge Communities. Management Science

Published online in Articles in Advance 12 Jun 2026

. https://doi.org/10.1287/mnsc.2024.04717

https://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–21
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

The AI Democratization Paradox: Evidence from Decentralized Knowledge Communities

Kai Zhu,[a,*] **Dylan Walker**[b]

[a] Bocconi University, 20136 Milan, Italy; [b] Chapman University, Orange, California 92866
*Corresponding author
Contact: kai.zhu@unibocconi.it, https://orcid.org/0000-0003-1954-0085 (KZ); dylan@chapman.edu, https://orcid.org/0000-0002-1354-4969 (DW)

Received: January 31, 2024
Revised: February 21, 2025; September 8, 2025; January 5, 2026
Accepted: March 23, 2026
Published Online in Articles in Advance: June 12, 2026

https://doi.org/10.1287/mnsc.2024.04717

Abstract. Does AI democratize knowledge production or amplify existing disparities? We investigate this tension by studying the deployment of neural machine translation across more than 100 Wikipedia language communities. Leveraging rich, fine-grained data and exogenous variation from a natural experiment, we uncover the "AI democratization paradox," where the technology simultaneously drives democratizing and concentrating forces. AI lowered barriers, leading to a substantial increase in content creation across diverse target languages without sacrificing quality or readership. However, the benefits were concentrated: well-resourced communities captured disproportionate gains—three to four times larger than mid-tier editions. Whereas editors actively leveraged AI to address representation gaps, translating female biographies at twice the expected rate, structural constraints still limited the impact in high-need areas. We conclude that technological solutions alone cannot overcome structural inequalities; AI's distributional impact is contingent on the interplay between technological capabilities and existing social structures.

History: Accepted by Anindya Ghose, information systems.

Funding: This work was supported by the Wikimedia Research & Technology Fund and European Union – NextGenerationEU funds, Component M4.C2, Investment 1.1, PRIN 2022 PNRR, CUP: J53D23015440001.
Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mnsc.2024.04717.

Keywords: artificial intelligence • machine translation • digital inequality • decentralized platforms • distributional effects • knowledge production • socio-technical systems

1. Introduction

Artificial intelligence (AI) is rapidly reshaping how knowledge is created, curated, and distributed globally (Agrawal et al. 2019, 2022; Goldfarb and Tucker 2019). This transformation is particularly pronounced within decentralized knowledge platforms, ranging from Wikipedia's multilingual encyclopedias to GitHub's collaborative coding repositories. These platforms operate as instances of commons-based peer production (Benkler 2002) within two-sided market structures (Rochet and Tirole 2003, Parker et al. 2016), relying on distributed contributors, often volunteers, to build vast information resources that constitute a modern digital commons (Ostrom 1990).

Within these unique ecosystems, AI tools, such as machine translation and large language models, promise to lower barriers to participation. However, their deployment raises a critical question that sits at the heart of contemporary debates on technology and inequality (Capraro et al. 2024): does AI democratize knowledge production by leveling the playing field, or does it amplify existing inequalities by disproportionately benefiting already advantaged actors?

This debate centers on two competing perspectives. The *technological optimism* view posits that AI acts as a democratizing force by reducing technical barriers, such as the need for multilingual proficiency, thereby enabling broader participation (Jensen 2007; Brynjolfsson et al. 2019, 2025). Conversely, the *structural reproduction* view argues that technology often reinforces existing disparities, as participants with greater complementary resources are better positioned to capture the benefits of new tools (Cohen and Levinthal 1990, Forman et al. 2012). Understanding which force dominates, and under what conditions, is crucial for designing equitable digital platforms and anticipating AI's societal impact.

We investigate this tension by examining the impact of AI-powered machine translation (MT) on translation-mediated content creation across Wikipedia's global platform. Our study leverages a unique natural experiment: the integration of Google Translate's advanced neural MT system into Wikipedia's Content Translation tool in January 2019. This setting provides an ideal laboratory. With over 100 language editions affected simultaneously, we can observe how the same technological shock interacts with diverse community structures, ranging from the massive English edition to smaller, resource-constrained languages, each governed by its own set of community-enforced norms and practices.

Employing a difference-in-differences strategy that exploits variation in the timing and coverage of Google Translate's rollout, we uncover what we term the *AI democratization paradox*. We find that advanced machine translation simultaneously unleashes both democratizing and concentrating forces.

On the one hand, the technology substantially democratizes content creation. We observe a remarkable 139% increase in translation-mediated content creation, generating an estimated 12.3 million additional pageviews over two years. Importantly, these gains are distributed across a diverse array of target languages, extending knowledge access to speakers of previously underserved linguistic communities worldwide. Beyond this quantitative expansion, editors actively leverage the efficient gains to address systemic representation gaps—female biographies, for instance, are translated at twice the rate expected based on existing content distributions.

On the other hand, concentrating forces significantly counteract these democratizing gains. The benefits flow disproportionately to well-resourced language communities, with the largest editions experiencing gains three to four times larger than mid-tier communities and over 10 times larger than the smallest editions. English Wikipedia's dominance as a source language intensifies from 68% to 81%, whereas regions with the greatest need, such as sub-Saharan Africa, see minimal impact.

Our study reconciles these paradoxical outcomes by demonstrating that AI's impact is critically shaped by the *interaction between technological capabilities and existing social structures*. Machine translation operates through an "expertise unbundling" mechanism, decoupling content knowledge from the need for specific language skills. This lowers the barrier to entry. However, realizing the benefits of this unbundling requires complementary resources. Our distributional analyses reveal that communities with greater editorial capacity, larger existing knowledge bases, and established reader networks are better equipped to leverage the technology, leading to the concentration of benefits. Technological capability alone cannot overcome structural inequalities.

This research makes two primary contributions. First, we quantify and contextualize the socioeconomic significance of AI-powered translation in a major commons-based peer production environment. By dramatically accelerating translation efforts—compressing what would have taken decades of manual labor into just two years—AI tools facilitated an estimated 12.3 million additional pageviews during our study period, a substantial expansion in access to knowledge. Second, we move beyond "AI increases productivity" narratives by providing novel, large-scale empirical evidence of how socio-technical systems shape AI's distributional impact. By foregrounding the AI democratization paradox, we offer a conceptual framework that reconciles competing views on technology's role in digital inequality.

Finally, our findings offer actionable insights for platform governance and AI deployment strategy. We demonstrate that providing equal access to AI tools is necessary but insufficient to ensure equitable outcomes. To realize AI's democratizing potential, platforms must implement complementary interventions, such as targeted editorial support, capacity building, and incentive structures, that address the resource constraints of disadvantaged communities. Organizations deploying AI in decentralized systems should adopt a distributional lens, anticipating not just aggregate productivity gains but also where technology will alleviate versus amplify existing disparities.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature and develops our theoretical framework. Section 3 describes the empirical setting, and Section 4 details the data used in our study. Section 5 outlines our empirical strategy. Section 6 presents the aggregate impact of AI on content production and consumption, whereas Section 7 analyzes the distributional effects across language communities and content representation, revealing the AI democratization paradox. Finally, Section 8 concludes with a discussion of the theoretical and managerial implications.

2. Theoretical Framework

This study addresses the intersection of AI adoption, digital platforms, and inequality. To articulate the central theoretical tension, we first define the economic and organizational context of our study—a decentralized knowledge platform—then examine competing perspectives on how AI deployment should affect such systems, and finally clarify our contribution.

2.1. Economic Organization of Decentralized Knowledge Platforms

The rapid advancement of AI is reshaping the digital economy (Agrawal et al. 2019, 2022). Whereas much

research focuses on AI adoption within firms, decentralized knowledge platforms present a distinct and theoretically important context. These platforms, including Wikipedia, GitHub, and Reddit, represent a third mode of economic organization, separate from traditional firms and markets, known as commons-based peer production (Benkler 2002). In these systems, large numbers of geographically dispersed individuals, driven by a diverse set of intrinsic and social motivations rather than direct financial compensation, work cooperatively to generate vast repositories of public information goods.

From a structural perspective, these platforms function as two-sided markets (or multisided markets), creating value by facilitating interactions between distinct groups of users—in Wikipedia's case, content contributors (editors) and content consumers (readers) (Rochet and Tirole 2003, Parker et al. 2016). The platform's operator (the Wikimedia Foundation) does not produce the content but provides the infrastructure and governance framework that enables the community to thrive (Teh 2022).

The governance of these platforms is also unique. Rather than relying on managerial hierarchy (as in a firm) or explicit contracts (as in a market), they are often self-governed by complex systems of community-developed rules, norms, and enforcement mechanisms. This model of collective action to manage a shared resource closely mirrors the principles for successful commons governance identified by Ostrom (1990). This socio-technical environment, characterized by volunteer labor, nonmarket incentives, and community-based governance, creates complex and unpredictable dynamics (Zhang and Zhu 2011, Faraj et al. 2016, Zhu et al. 2020). The deployment of AI in such a context raises a fundamental question: will it empower the collective and level the playing field, or will it exacerbate the structural inequalities already present in the system?

2.2. Democratizing Potential of AI: A Technological Optimism Perspective

The technological optimism perspective suggests that AI can act as a democratizing force. By reducing technical barriers and transaction costs, such as the need for multilingual proficiency, AI enables broader participation. This view, articulated in the theory of democratizing innovation, posits that as the costs of design and production tools fall, the locus of innovation can shift from centralized producers to individual users, who can now create what they need for themselves (Von Hippel 2006).

In multilingual environments such as Wikipedia, language barriers represent a significant impediment to knowledge sharing, creating what scholars term the "digital language divide." Machine translation, powered by recent advances in neural networks (Wu et al. 2016, Vaswani et al. 2017), directly addresses this constraint by lowering the transaction costs of cross-linguistic communication. Prior research provides strong causal evidence for this effect, showing that the introduction of high-quality machine translation on an e-commerce platform significantly increased international trade by making it easier for buyers and sellers to understand each other (Brynjolfsson et al. 2019).

In the context of knowledge production, this leads to what we term an expertise unbundling mechanism. Traditionally, creating a Wikipedia article in a new language required both domain expertise (knowledge of the topic) and specific production skills (fluency in the target language). Advanced machine translation decouples these requirements. It allows a contributor with deep knowledge of a topic in one language to create a high-quality draft in another language without being fluent in it, leaving the final refinement to a native-speaking editor. Theoretically, this unbundling should flatten hierarchies and enable wider participation, particularly from under-resourced language communities seeking to import knowledge from larger ones. This process should, in theory, accelerate the growth of smaller encyclopedias by tapping into the cumulative growth dynamics observed on the platform, where a larger base of existing content attracts even more contributions (Aaltonen and Seiler 2016).

We test this democratizing prediction through our aggregate effects analysis (Section 6) and content representation analysis (Section 7.2), examining whether AI expands overall production and enables editors to address representation gaps.

2.3. Concentrating Effects of AI: A Structural Reproduction Perspective

Conversely, a substantial body of literature on socio-technical systems and digital inequality cautions that technology adoption often reinforces, rather than reduces, existing disparities. This structural reproduction view argues that the benefits of new technologies are not uniformly distributed but are instead mediated by existing social structures and the availability of complementary resources (Wessels 2013). The features of a technology do not operate in a vacuum; their impact depends on how users can leverage them within specific social contexts.

This perspective suggests that communities already possessing advantages, such as a large base of active editors, a substantial existing knowledge repository to translate from, and a large, engaged readership, are better positioned to adopt and exploit new AI tools. For example, a language community with many active editors can more effectively review, correct, and integrate the surge of machine-translated drafts, whereas a community with few editors may be overwhelmed. This differential capacity can lead to a concentration

of benefits among the already advantaged, a dynamic that parallels findings from labor economics suggesting AI's gains will flow disproportionately to highly skilled workers who can complement the technology (Cazzaniga et al. 2024).

Indeed, extensive research has documented persistent structural inequalities within Wikipedia, despite its open and collaborative mission. Significant gaps exist in content representation across gender, geography, and language dimensions (Graham et al. 2014, Wagner et al. 2016, Miquel-Ribé and Laniado 2020). From a structural reproduction viewpoint, without carefully designed interventions, AI deployment risks exacerbating these gaps by primarily accelerating content creation in the very communities that are already dominant. This aligns with a broader historical perspective, which posits that power relations and institutional capacity, not technology alone, are the ultimate determinants of how the gains from innovation are distributed across society (Cohen and Levinthal 1990, Scheidel 2024).

We test this concentration prediction through heterogeneity analyses (Section 7.1), examining whether communities with greater complementary resources—larger editor bases, more existing content, larger readerships—capture disproportionate gains.

2.4. AI Democratization Paradox and Our Contribution

These competing perspectives—technological optimism and structural reproduction—create a critical theoretical tension regarding AI's impact on decentralized knowledge production. Our contribution is the articulation and empirical validation of the *AI democratization paradox*: the finding that a single technological intervention simultaneously drives both democratization and concentration.

The paradox arises because AI lowers barriers by decoupling content creation from language-specific skills. It expands participation (democratizing) while simultaneously increasing the relative importance of complementary resources that remain unequally distributed (concentrating). Whether democratizing or concentrating forces dominate depends on which factor—language skills or complementary resources—constitutes the binding constraint in a given context. Democratizing effects are most visible at the *individual level*, where motivated agents can leverage AI to act on preferences previously blocked by language limitations. Concentrating effects dominate at the *community level*, where structural advantages—editorial capacity, existing knowledge bases, reader networks—compound to capture disproportionate gains.

This framework helps explain why simply providing equal access to a powerful tool may be insufficient to address digital inequalities. Our research moves beyond asking *if* AI increases productivity to ask *for whom* it increases productivity and *under what structural conditions*. By connecting theoretical predictions to specific empirical tests, we provide a systematic account of when AI democratizes versus when it amplifies existing disparities.

3. Setting

Wikipedia serves as an ideal empirical setting to examine the interaction between AI deployment and decentralized knowledge production. As the largest collaborative encyclopedia in the world, Wikipedia operates through a volunteer-driven model across more than 300 language editions, with each language community possessing vastly different levels of resources and organizational capacity. This heterogeneity enables us to study how a uniform technological shock—AI-powered translation—interacts with diverse existing social structures.

3.1. The Challenge: Cross-Linguistic Knowledge Gaps

Despite Wikipedia's global reach, knowledge availability varies sharply across languages. Content overlap between language editions is limited, meaning that readers of one language often lack access to information that is widely consumed in another.

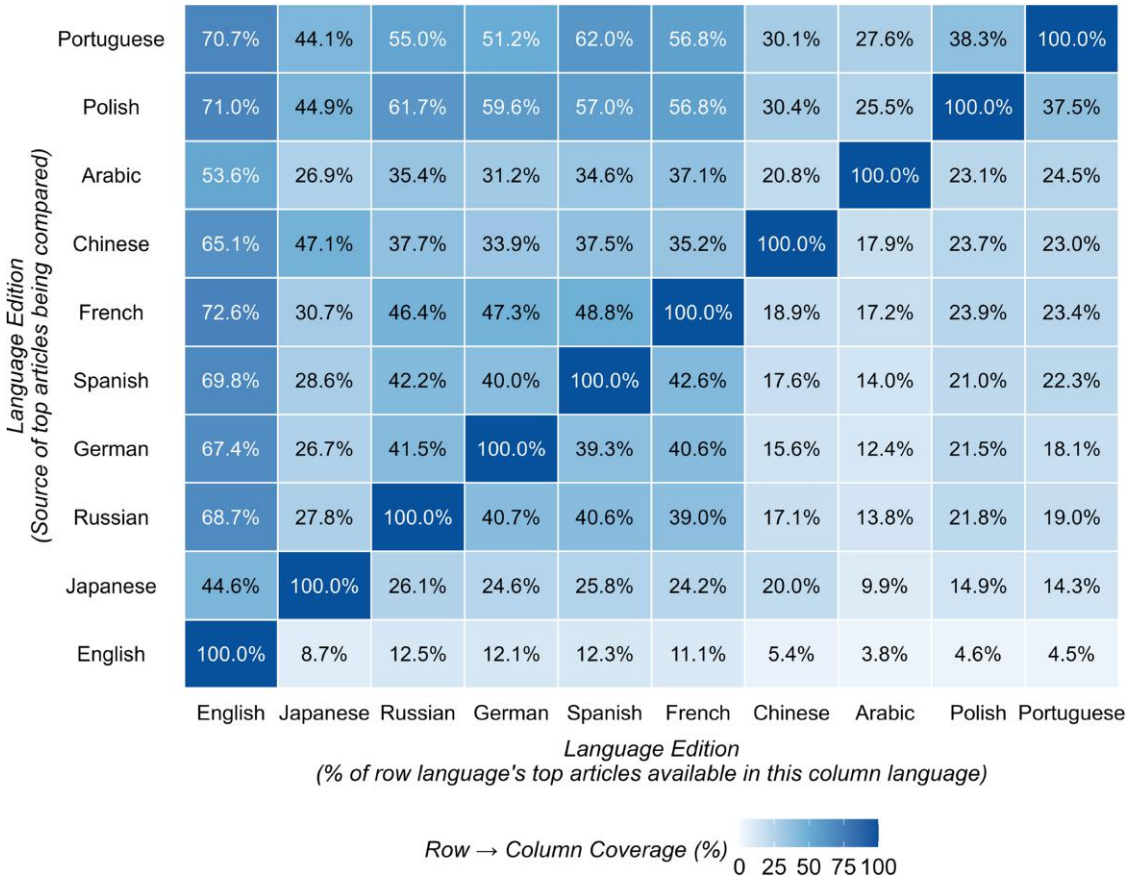
Figure 1 quantifies these differences using the top 5% most viewed articles in 2024 for the 10 largest language editions. Each cell reports *directional coverage*: the share of high-traffic articles in the row language that also exist in the column language. Two patterns stand out. First, coverage is generally low—pairwise overlap rarely exceeds 70% even among major languages. Second, coverage is asymmetric: for example, more than 60%–70% of Portuguese or Spanish high-traffic articles appear in English, whereas only about 5%–13% of English high-traffic articles appear in most other editions (see the bottom row). These patterns indicate substantial cross-linguistic gaps and suggest that much of Wikipedia's most demanded knowledge remains siloed by language.

Historically, creating an article in a new language required writing from scratch—a time-intensive process demanding both domain expertise and proficiency in the target language. To alleviate this constraint, the Wikimedia Foundation introduced the Content Translation system in 2015, which enables *translation-mediated content creation* and lowers the barriers to systematic knowledge transfer across language editions.

3.2. The Human-AI Collaboration Workflow

The Content Translation system implements a human-in-the-loop workflow that combines machine translation

Figure 1. (Color online) Large and Asymmetric Cross-Language Gaps in High-Traffic Wikipedia Content



Notes. Each cell shows the share (percentage) of the row language's top 5% most viewed 2024 articles that are also available in the column language (*row* → *column* coverage). Diagonal cells equal 100% by construction. Pairwise coverage is generally low and asymmetric: coverage seldom exceeds 70%, and English in particular exhibits low outward coverage (roughly 5%–13% into other editions) while covering a large share of other languages' high-traffic content.

capabilities with human editorial oversight. When editors select an article for translation, the system presents a dual-panel interface (Figure 2): the source article appears on one side, whereas a workspace for creating the translated version appears on the other. Importantly, the system automates technical aspects—preserving formatting, maintaining link structures, and transferring citations—allowing editors to focus on ensuring linguistic accuracy and cultural adaptation.

The system generates machine translations as initial drafts, which editors then review, revise, and refine before publication. This structured human-AI collaboration is essential for maintaining Wikipedia's quality standards while leveraging the efficiency gains from automation. The human-in-the-loop design distinguishes this setting from fully automated translation applications.

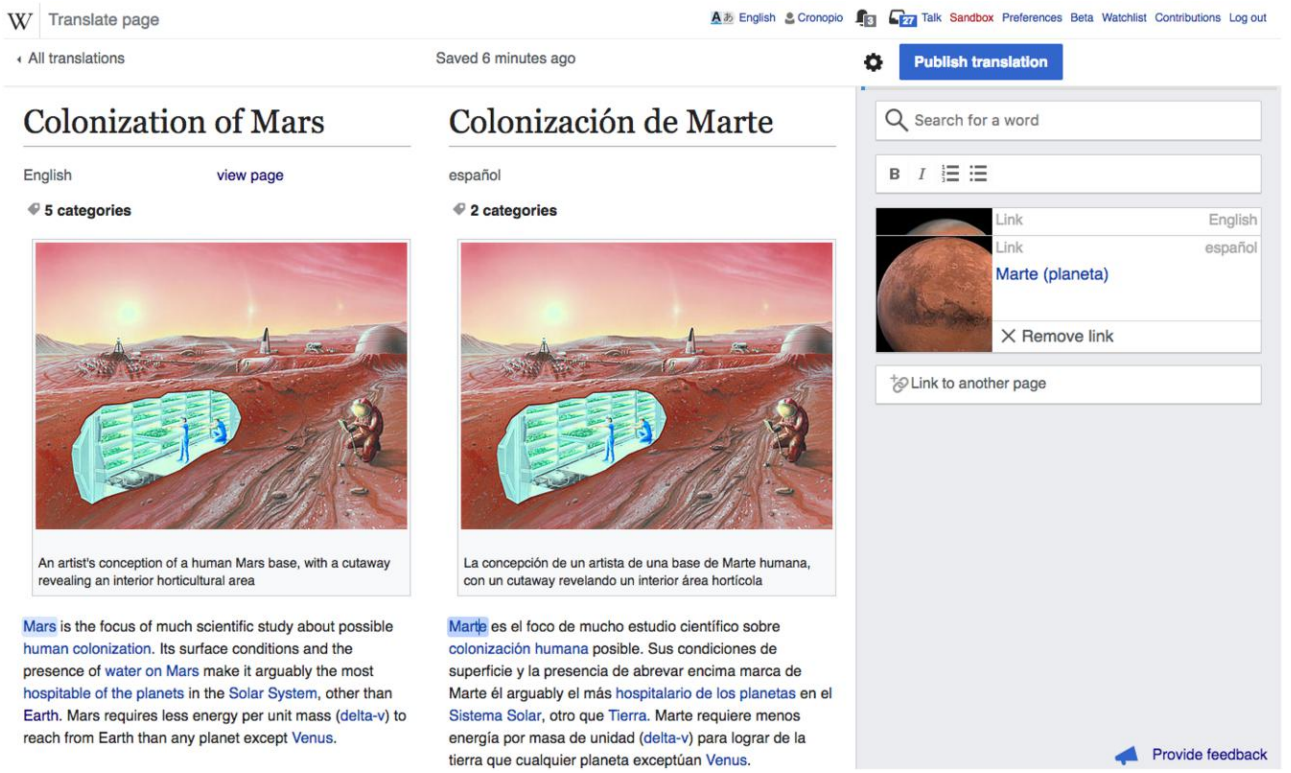
By late 2024, this system had facilitated the creation of over two million articles—a volume equivalent to creating three medium-sized Wikipedia language editions or roughly one-third the size of English Wikipedia. This substantial output demonstrates the system's transformative impact on cross-linguistic knowledge transfer.

3.3. The Technological Shock: Google Translate Integration

The Content Translation system initially relied on basic machine translation engines such as Apertium and Yandex, which offered limited accuracy and language coverage. The critical intervention we study occurred in January 2019, when the Wikimedia Foundation integrated Google Translate's advanced neural machine translation system (Figure 3).[1]

This integration represented a substantial technological shock to the Wikipedia ecosystem. Google Translate employs neural machine translation (NMT) architecture, which processes entire sentences holistically rather than word by word, capturing context and linguistic nuance far more effectively than previous statistical methods. Published evaluations demonstrated that this technology reduced translation errors

Figure 2. (Color online) The Human-AI Collaboration Workflow in Wikipedia's Content Translation Tool

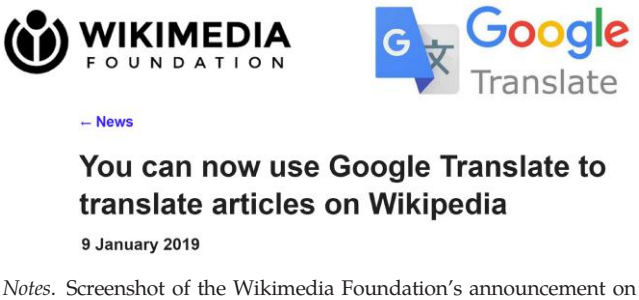


Notes. Screenshot of the Content Translation tool interface. The system employs a human-in-the-loop workflow, presenting the source article (left) alongside a workspace for the translated version (right). It generates machine translations as initial drafts, which editors then review, revise, and refine before publication. The system also automates technical aspects such as formatting and link preservation.

by 55%–85% across major language pairs (Wu et al. 2016), substantially lowering the cognitive burden for editors engaged in translation-mediated content creation.

This empirical setting offers several distinct advantages for our research. First, the integration provides a clear, plausibly exogenous technological shock affecting over 100 language communities simultaneously.

Figure 3. (Color online) The Technological Shock: Integration of Google Translate's Neural Machine Translation System



Notes. Screenshot of the Wikimedia Foundation's announcement on January 9, 2019, regarding the integration of Google Translate into the Content Translation tool. This integration introduced advanced neural machine translation (NMT) capabilities, representing a substantial, exogenous technological shock by significantly improving translation quality and lowering the cognitive burden for editors.

Second, Wikipedia's open architecture provides comprehensive data on content creation patterns, quality metrics, and user engagement. Finally, the platform's decentralized, volunteer-driven nature enables us to study how AI capabilities interact with existing community structures, offering key insights into the conditions under which technology democratizes or concentrates knowledge production.

4. Data

4.1. Data Sources and Sample Construction

Our empirical analysis draws on rich data from Wikipedia's decentralized ecosystem, enabling a detailed examination of both the aggregate and distributional effects of AI deployment on translation-mediated content creation.

4.1.1. Primary Data: Translation Records. Our core data set comprises complete translation records from Wikipedia's Content Translation system, which logs every article translation facilitated through the tool. Each record contains detailed metadata, including source and target languages, article titles, publication timestamps, and the specific machine translation engine used (see Section A.1 in the Online Appendix

for technical details). Our sample spans January 2017 to December 2020, providing two years of observations before and after the primary integration of Google Translate in January 2019.

4.1.2. Complementary Data Sources. We enrich our primary data set with multidimensional metrics at the article, language, and editor levels to capture the full scope of the intervention's impact.

Article-level metrics: To measure reader engagement, we use pageview statistics. To assess content quality, we employ two complementary signals: the Wikimedia Foundation's language-agnostic structural quality scores and community deletion logs that track articles removed for failing to meet standards. To analyze representation patterns, we leverage Wikidata's structured metadata providing demographic information (gender for biographical articles, geographic location for place-based content) and employ language-agnostic topic classifications (Johnson et al. 2021) to systematically examine content gaps across domains.

Language-level characteristics: To understand how existing social structures mediate AI's impact, we compile metrics capturing each language community's complementary resources. These include (1) knowledge base size, measured by total existing articles; (2) editorial capacity, captured through active editor counts; and (3) readership network size, indicated by total in the language edition.

4.1.3. Sample Construction. We applied several filtering criteria to construct our analysis sample. We focus our analysis on target languages, including only those with at least 20 translations during the sample period. We exclude editions exhibiting abnormal translation spikes suggestive of bot activity or coordinated campaigns[2] and remove English Wikipedia as a translation destination because of its unique content policies restricting translations from other languages.

This process yields a substantial data set of 540,488 translated articles—a volume equivalent to the entire content of a small- to medium-sized Wikipedia language edition. For panel analyses, we aggregated these articles to the language-month level, creating a balanced panel data set of 132 languages observed over 48 months from January 2017 to December 2020.

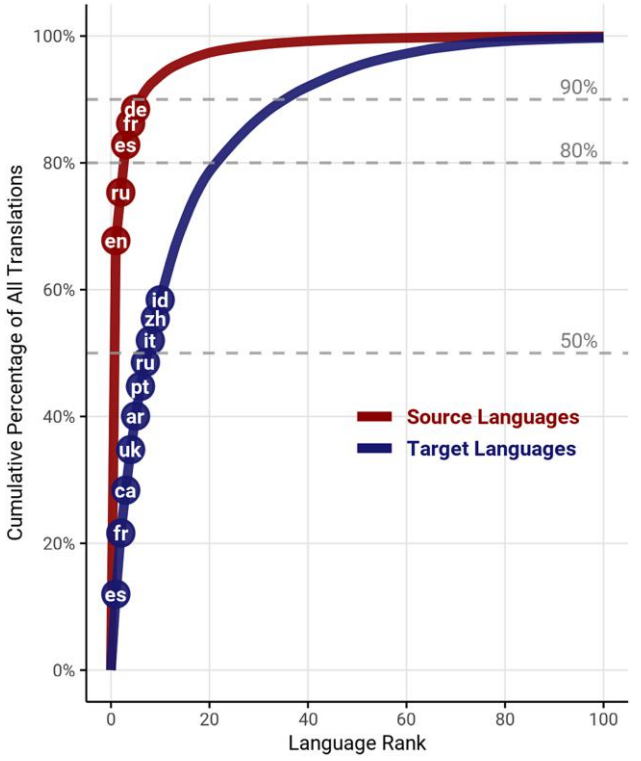
4.2. Pre-intervention Patterns and Structural Asymmetries

The baseline distribution of translation activity reveals structural asymmetries that existed before the AI intervention, providing key context for interpreting our subsequent findings. These pre-existing patterns highlight a centralized knowledge-sharing network characterized by structural asymmetries.

Figure 4 illustrates the stark concentration of translation activity. English Wikipedia dominates as the information source for 68% of all translations, establishing its role as the primary global knowledge hub. The top five source languages—English, Russian, Spanish, French, and German—collectively account for nearly 90% of translated content. Whereas target languages show greater diversity, concentration remains significant: the top 10 recipient languages receive 58% of all translations, rising to 95% for the top 50 languages. These patterns reveal that, whereas translation extends knowledge beyond its original sources, the benefits of knowledge transfer are heavily concentrated in a relatively small subset of Wikipedia's 300+ language editions.

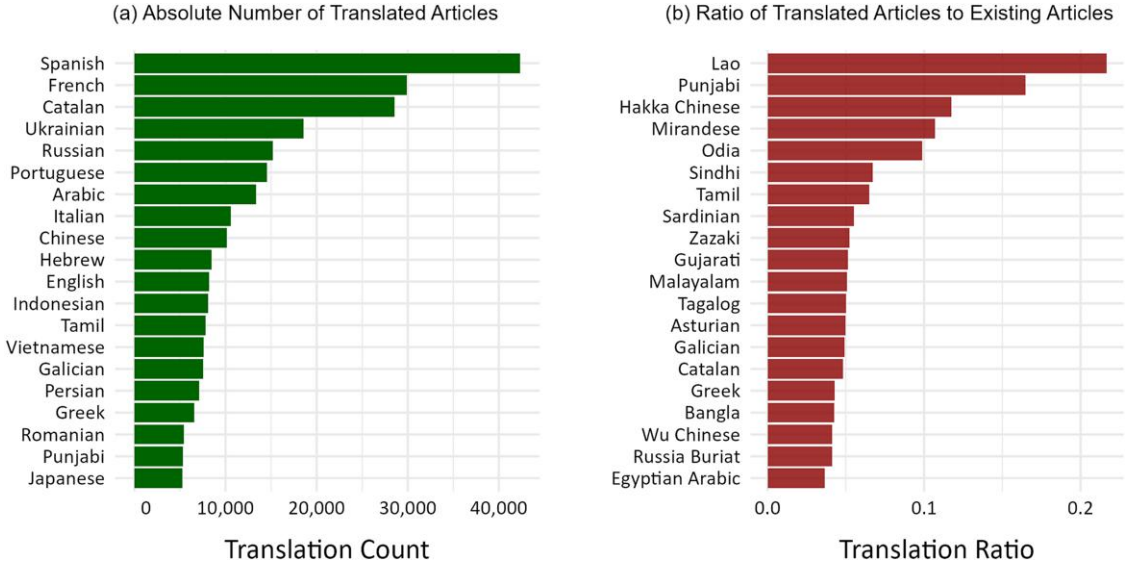
This concentration masks a *dual inequality* in how different communities depend on translation, as revealed in Figure 5. Figure 5(a) shows that large, well-resourced communities dominate in absolute volume—Spanish leads with approximately 45,000 translated articles, followed by French (35,000) and Catalan (30,000). These communities possess the editorial capacity to absorb and integrate substantial translated content.

Figure 4. (Color online) Pre-intervention Centralization of the Global Knowledge-Sharing Network (2017–2018)



Notes. Cumulative distribution of translation activity across source and target languages before the Google Translate intervention. The source languages (red) are highly concentrated: English dominates as the source for 68% of all translations, and the top five languages account for nearly 90%. Target languages (blue) show more diversity, yet the top 10 recipient languages still receive 58% of all translations.

Figure 5. (Color online) The Dual Inequality in Pre-intervention Translation Dependency (2017–2018)



Notes. This figure illustrates the structural asymmetries in translation reliance across Wikipedia language editions before the Google Translate intervention. (a) The absolute volume of translated articles by target language, indicating that large, well-resourced communities (e.g., Spanish, French) dominate in total volume. (b) The ratio of translated articles to the total existing articles for each language, revealing that smaller, resource-constrained communities (e.g., Lao, Punjabi) rely much more heavily on translation (15%–20% of total content) for their overall growth.

However, Figure 5(b) reveals a strikingly different pattern when examining relative dependency. Smaller, resource-constrained communities rely most heavily on translation as a proportion of their total content. In language editions such as Lao, Punjabi, and Hakka Chinese, translated articles constitute 15%–20% of the entire encyclopedia, suggesting that for these smaller editions, translation functions not merely as a supplement but as a critical engine of growth.

These baseline patterns establish the asymmetric structure that predated AI intervention: large language editions dominate absolute translation volumes, whereas smaller editions depend more critically on translation for their development. This dual inequality provides essential context for understanding how a technological shock such as Google Translate deployment interacts with existing disparities to produce the democratization paradox we identify.

4.3. Descriptive Evidence: Technology Adoption and Workflow Shift

Initial descriptive analysis reveals dramatic behavioral shifts following Google Translate's integration, providing preliminary evidence of the technology's transformative impact. Figure 6 documents these changes from two complementary perspectives.

Figure 6(a) tracks the usage patterns of different machine translation engines over time. Before January 2019 (marked by the vertical line), Yandex and Apertium maintained stable but modest usage levels. Following integration, Google Translate exhibits rapid adoption, simultaneously displacing existing engines and dramatically expanding overall machine-assisted translation volume. This pattern suggests that Google Translate did not merely substitute for existing tools but fundamentally expanded the scope of AI-assisted content creation.

Figure 6(b) reveals the corresponding transformation in editor workflows. The proportion of manual translations—articles created without any machine assistance—dropped precipitously from approximately 20% before integration to single digits afterward. This near-complete transition to AI-assisted workflows indicates that the improved technology became the de facto standard for cross-linguistic content creation.

Together, these patterns provide suggestive evidence of a fundamental transformation in content creation processes, motivating our formal causal analysis of the technology's aggregate and distributional effects.

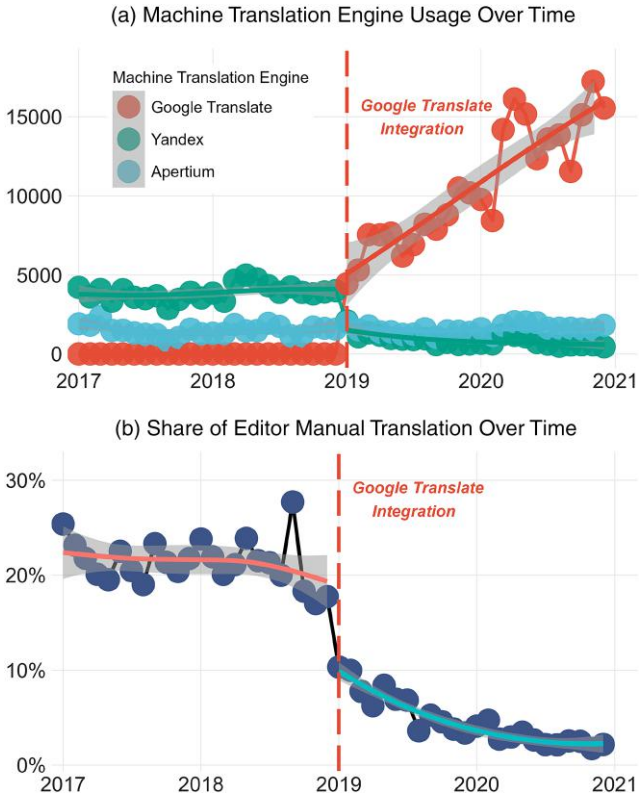
5. Empirical Strategy

We exploit Google Translate's integration into Wikipedia's Content Translation system as a quasi-experimental shock to identify the causal impact of advanced AI on knowledge production. Our identification strategy leverages variation in both the timing and coverage of this technological deployment.

5.1. Identification Strategy and Treatment Definition

The integration of Google Translate provides a plausibly exogenous source of variation for causal identification. The deployment resulted from a top-down partnership

Figure 6. (Color online) Rapid Adoption of Advanced Machine Translation and the Transformation of Translation Workflows (2017–2020)



Notes. Behavioral shifts following the integration of Google Translate (marked by the vertical red line in January 2019). (a) Track of the monthly usage of different machine translation engines, showing the rapid adoption of Google Translate, which both displaced existing engines (Yandex, Apertium) and expanded overall volume. (b) The proportion of articles translated manually (without machine assistance), which dropped sharply from approximately 20% to single digits, indicating a near-complete transition to AI-assisted workflows.

between the Wikimedia Foundation and Google, making the timing independent of individual language communities' activities or characteristics. Moreover, treatment assignment was determined by Google's pre-existing language support rather than Wikipedia-specific factors, creating quasi-random variation in AI availability.

Within our sample of 132 Wikipedia language editions, the deployment occurred in waves (Table 1). The primary wave in January 2019 simultaneously treated 78 languages, followed by staggered additions of 12 languages through August 2020. The remaining 42 languages, unsupported by Google Translate during our sample period, serve as controls. This staggered structure, while creating some complexity, provides rich variation for identification, with the vast majority of treated languages (78 of 90) receiving simultaneous access.

This implementation pattern enables a difference-in-differences research design with our language-month

Table 1. Staggered Rollout of Google Translate Across Wikipedia Language Editions

Implementation status	Deployment period	Languages
Completed implementation (treated)		
Initial major rollout	January 2019	78
Staged rollout	June 2019–August 2020	12
Pending implementation (control)	—	42
Total in sample		132

Notes. Summary of the Google Translate deployment timeline across the 132 languages in our analysis sample. The deployment occurred in waves, providing variation in timing and coverage used for the difference-in-differences identification strategy. The 42 languages unsupported during the sample period serve as controls.

panel data. Our identification relies on the parallel trends assumption: absent the intervention, treated and control languages would have followed similar trajectories in translation activity. We validate this assumption through multiple pre-trend tests and placebo analyses presented below.

5.2. Primary Estimation: Counterfactual Imputation Approach

Given the staggered rollout and potential for heterogeneous treatment effects, we employ contemporary difference-in-differences methods that address recent econometric concerns about traditional two-way fixed effects (TWFE) estimators. Our primary approach uses counterfactual imputation, with robustness checks using multiple alternative estimators, including stacked difference in differences (DiD) (Cengiz et al. 2019), interaction-weighted estimation (Sun and Abraham 2021), and doubly robust methods (Callaway and Sant'Anna 2021). All approaches yield consistent results (detailed in Online Appendix C).

We adopt counterfactual imputation as our primary method for three reasons. First, causal inference fundamentally addresses a missing data problem (Holland 1986)—we observe treated units' outcomes under treatment but not their counterfactual outcomes without treatment. Directly imputing these counterfactual outcomes provides a natural and transparent framework for estimating treatment effects. Second, this approach accommodates both staggered rollout timing and treatment effect heterogeneity across units and time periods. Third, it avoids the negative weighting problems that have been highlighted in recent applied econometrics literature regarding traditional TWFE models (De Chaisemartin and d'Haultfoeuille 2020, Goodman-Bacon 2021).

Specifically, we implement the matrix completion estimator within the counterfactual imputation framework (Athey et al. 2021).[3] The matrix completion method constructs a panel data matrix with language editions as rows, time periods as columns, and translation volumes

as entries. The method leverages patterns across both dimensions—using information from control units and pre-treatment periods—to impute counterfactual outcomes for treated observations. This approach is particularly well suited to our setting, as it flexibly accommodates the heterogeneous responses we expect across Wikipedia's diverse language communities.

5.3. Alternative Identification and Estimation

To strengthen causal inference and address potential concerns, we implement two complementary identification strategies that rely on different sources of variation and identifying assumptions.

5.3.1. Time-Shift Difference-in-Differences Design. As an alternative approach, we implement a time-shift difference-in-differences design focusing exclusively on the 78 languages treated in January 2019. We compare their translation activity in July 2018–June 2019 (post-treatment) against July 2017–June 2018 (pre-treatment), using the same languages' earlier activities as the counterfactual.

This within-language comparison alleviates concerns about unobserved heterogeneity affecting growth trajectories differently between different language groups. The limitation of this approach is the necessarily shorter time window, as we must reserve the earlier period as a control.

5.3.2. Blocked Difference-in-Differences Design. For our heterogeneity analyses examining how effects vary across language characteristics and content types, we employ a simplified blocked difference-in-differences design. This approach focuses on the 78 languages treated simultaneously in January 2019, excluding those with staggered timing. This creates a clean setting with single treatment timing, enabling transparent estimation using classical two-way fixed effects:

$$Y_{it} = \alpha \times GT_{it} + \boldsymbol{\beta} \times GT_{it} \times \boldsymbol{X_i} + \gamma_i + \delta_t + \epsilon_{it},$$

where Y_{it} represents translation volume for language i in month t, GT_{it} indicates Google Translate availability (one for treated languages post-January 2019, zero otherwise), $\boldsymbol{X_i}$ captures time-invariant language characteristics (e.g., editorial capacity, knowledge base size), and γ_i and δ_t are language and time fixed effects. The coefficient vector $\boldsymbol{\beta}$ identifies heterogeneous treatment effects—our primary interest in understanding the distributional impact of AI deployment.

This multipronged empirical strategy, combining different estimators and identification approaches, ensures our findings are robust to various econometric considerations and provides convincing evidence on both the aggregate and distributional effects of AI-powered translation.

6. Aggregate Impact: The Democratizing Potential of AI

We begin by examining the aggregate impact of Google Translate integration on translation-mediated content creation. This analysis tests the technological optimism perspective: whether advanced AI successfully lowers barriers to knowledge production through expertise unbundling and whether this increased output maintains quality standards and reader value. We employ our counterfactual imputation framework for language-level analysis and comparative analysis over time for article-level quality metrics.

6.1. The Productivity Surge

We first analyze the impact of Google Translate integration on the volume of translation-mediated content creation. This analysis tests whether advanced AI effectively decouples language skills from content knowledge, thereby increasing output.

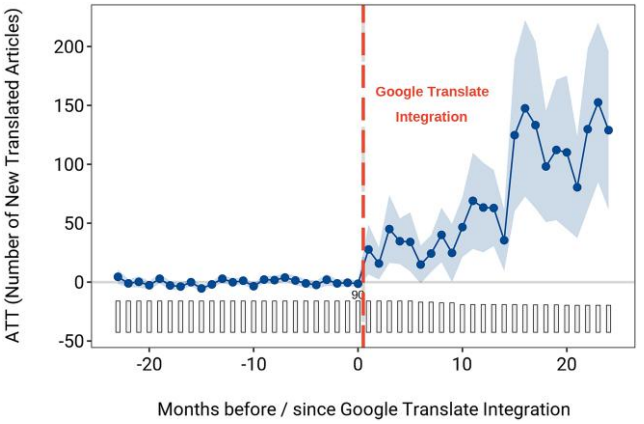
6.1.1. Main Results. The event study results, estimated using the counterfactual imputation method, are presented in Figure 7. We observe parallel trends prior to the intervention, followed by an immediate and sustained increase in translation activity post-integration.

Our difference-in-differences estimates show that Google Translate's integration increased translation-mediated article creation by 71.2 articles per language edition per month (average treatment effect on the treated (ATT); $p < 0.01$). In the steady-state phase—the final six months of the post-integration period—the monthly increase rises to 119 translated articles, a 139% gain over the pre-integration average of 85.4 articles per month. Together, these magnitudes highlight the transformative "expertise unbundling" mechanism: by sharply lowering the linguistic skill required for translation, machine translation democratized cross-linguistic content creation.

6.1.2. Diagnostic Tests. In line with best practices in difference-in-differences analysis, we conduct two diagnostic tests (Figure 8) to strengthen our causal inference. First, an equivalence test confirms the absence of differential pre-trends ($p = 0.863$), validating our parallel trends assumption. Second, a placebo test artificially shifting the intervention six months earlier to August 2018 yields no significant effect, confirming that our model correctly identifies null effects when no true intervention has occurred. Together, these diagnostics provide strong support for our causal interpretation of the observed increases in translation activity post-integration.

6.1.3. Further Robustness. Our findings are robust across a range of contemporary difference-in-differences estimators, including Callaway and Sant'Anna's doubly

Figure 7. (Color online) The Technology-Driven Productivity Surge: A Significant Increase in Translation-Mediated Content Creation



Notes. Event study plot showing the impact of Google Translate integration ($t = 0$) on monthly translation volume, estimated using the counterfactual imputation method. The plot shows parallel trends prior to the intervention, followed by an immediate and sustained increase in activity. The estimated average treatment effect on the treated (ATT) is an increase of 71.2 additional articles per language/month from pre-integration levels. The blue line represents point estimates, and the shaded area indicates 95% confidence intervals.

robust estimator, Sun and Abraham's interaction-weighted approach, Cengiz et al.'s stacked DiD design, and traditional two-way fixed effects models. As an additional validation, our within-language time-shift design—which uses the same languages' earlier activities as the controls—produces highly consistent results (ATT = 79.1 articles/month). Together, this robustness across methodologies with varying identifying assumptions, weighting schemes, and control group definitions strengthens confidence in the causal validity of our estimates. Detailed results are reported in Online Appendix C.

6.2. Maintaining Quality at Scale

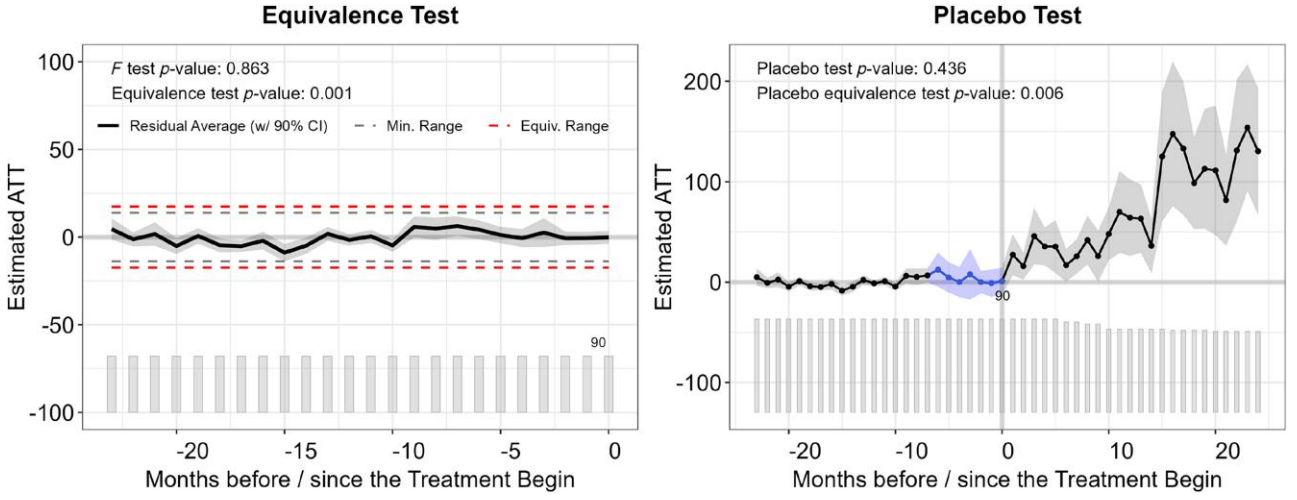
A critical concern with AI-assisted content creation is whether rapid scaling compromises quality. We address this through two complementary quality assessments: community validation through deletion processes and algorithmic structural quality evaluation.

6.2.1. Community Quality Standards. Wikipedia employs a collaborative deletion process to remove articles that fail to meet quality standards. The deletion ratio serves as a key indicator of community validation and represents the primary quality signal monitored by the Wikimedia Foundation in managing the Content Translation system.[4]

Figure 9(a) reveals a surprising finding: deletion rates actually *declined* following Google Translate integration. Rather than degrading quality through automation, the improved translation technology, operating within Wikipedia's human-in-the-loop framework, actually enhanced editors' ability to create content that meets community quality standards.

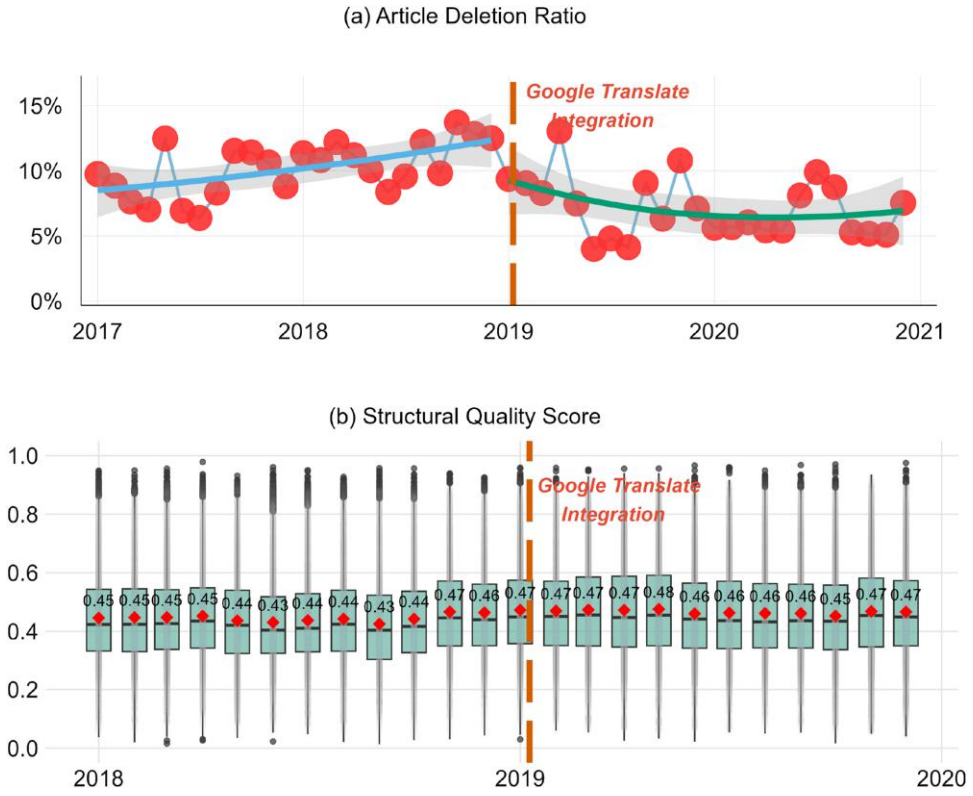
6.2.2. Objective Quality Assessment. We complement community-based validation with the Wikimedia Foundation's language-agnostic quality prediction model, which evaluates articles based on structural features such as reference density, section organization, article length, and multimedia elements. This algorithmic assessment generates a structural quality score for each Wikipedia article across different languages.

Figure 8. (Color online) Validation of the Causal Identification Strategy: Pre-trend and Placebo Tests



Notes. Diagnostic tests validating the difference-in-differences strategy. The left panel presents the equivalence test, which confirms the absence of differential pre-trends ($p = 0.863$), supporting the parallel trends assumption. The right panel presents the placebo test, which artificially shifts the intervention six months earlier (August 2018; blue region) and correctly identifies no significant effect.

Figure 9. (Color online) Content Quality Was Maintained Amid the Technology-Driven Productivity Surge



Notes. Analysis of content quality before and after Google Translate integration (January 2019). (a) Tracks the community validation metric (deletion ratio: proportion of translated articles subsequently deleted). The locally estimated scatterplot smoothing (LOESS) curves show that deletion rates declined post-integration. (b) The distribution of algorithmic structural quality scores (based on features like reference density and section organization). The improvement in median scores indicates that structural quality was maintained or slightly enhanced following Google Translate deployment ($t = 0$).

Figure 9(b) shows the distribution of structural quality scores for articles created before and after AI integration. We observe a modest but statistically significant improvement in structural quality following integration, with an average increase of 0.02 units ($p < 0.01$; detailed results in Section E.1 in the Online Appendix). This suggests that AI assistance contributed to the creation of better-structured and documented content.

Together, these complementary quality measures demonstrate that the dramatic surge in AI-assisted content creation did not come at the expense of content quality. Both community validation through deletion processes and objective structural assessments confirm that advanced machine translation, when integrated within collaborative human-AI workflows, maintained content standards while enabling production at substantially greater scale.

6.3. Sustained Reader Engagement

Beyond production metrics, we next examine the downstream impact on knowledge consumption to assess whether the surge in AI-assisted content creation translated into meaningful reader engagement and knowledge access.

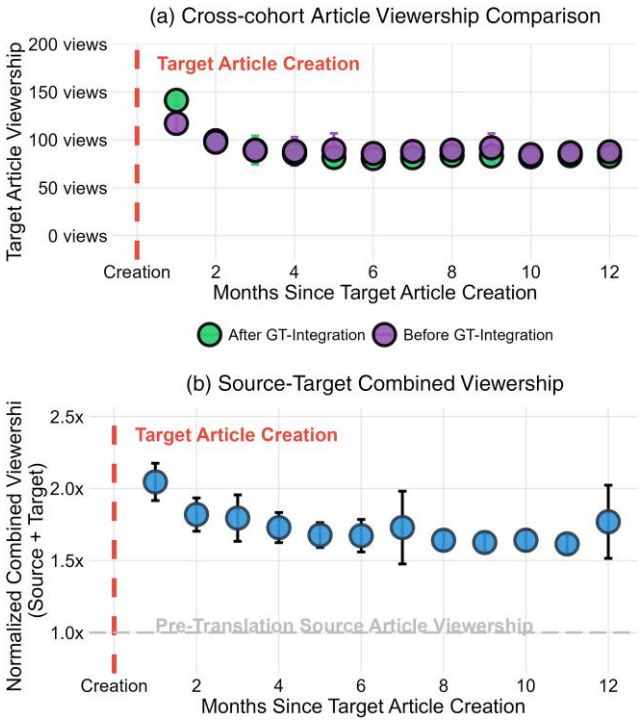
6.3.1. Content Utility.

We analyze article pageviews to assess the utility of translated content from the readers' perspective. Figure 10(a) presents a *cross-cohort comparison* of pageview trajectories, contrasting two distinct groups of articles: those translated *before* Google Translate integration (pre-2019 cohort) and those translated *after* integration (post-2019 cohort). For each cohort, we track monthly pageviews starting from the article's creation month (normalized as $t = 0$).

Both cohorts exhibit remarkably consistent patterns: an initial spike of 140–150 views in the creation month, followed by stabilization at 80–90 monthly views. The two trajectory lines are statistically indistinguishable—we observe *no significant difference* in readership patterns between the pre- and post-integration cohorts.

This cross-cohort equivalence represents a key finding. Despite the 139% surge in translation volume following AI integration, the per-article reader engagement remained constant. Articles produced under the AI-assisted regime attract the same

Figure 10. (Color online) Sustained Reader Engagement and the Expansion of Knowledge Access



Notes. Analysis of reader engagement (pageviews) with AI-assisted content. (a) Comparison of the monthly pageview trajectories for articles created before (purple) and after (green) Google Translate integration ($t = 0$ is the creation month). Both cohorts show consistent engagement, stabilizing at 80–90 monthly views, indicating maintained content utility despite the volume surge. (b) Analysis of the combined readership of source-target article pairs, normalized against the pre-translation source readership (baseline = 1.0). The combined readership increases to approximately 1.6 times the baseline, demonstrating genuine consumption expansion rather than substitution.

readership as those created through prior methods. This pattern contradicts concerns that AI-assisted content might be lower quality or less useful to readers, instead validating the effectiveness of the human-AI collaborative workflow in maintaining content utility at scale.

6.3.2. Expansion vs. Substitution. An interesting question is whether translated content expands overall knowledge consumption or substitutes readership from source languages. Either outcome would be valuable: even pure substitution would enable readers to access content in their preferred language. To explore this, we analyze the combined readership of source-target article pairs, normalizing against the pre-translation source readership baseline (set to 1.0).

Figure 10(b) reveals that following translation, combined readership increases to approximately 1.6 times the baseline and persists at this elevated level. This demonstrates genuine consumption expansion: the availability of content in target languages attracts new readership rather than diverting existing readers from source articles. This finding indicates that AI-assisted translation creates net value by making knowledge accessible to previously underserved linguistic communities.[5]

These results offer revealed-preference evidence that supply is meeting genuine demand. If reader interest were absent, engagement would be minimal; instead, we observe sustained viewership (80–90 monthly views; 75th–80th percentile) and a 1.6× expansion in combined readership. This indicates that AI lowered barriers to unlock content that readers actively value, rather than merely inflating supply.[6]

6.4. Economic and Social Significance

Whereas 80–90 monthly views per article might appear modest in isolation, the aggregate impact is substantial and economically significant in the context of global knowledge access.

The scale of aggregate knowledge consumption is remarkable. Our estimates imply approximately *12.3 million additional pageviews* over our two-year study period—knowledge consumption that would not have occurred without AI assistance.[7] This equals the total traffic of a midsized Wikipedia edition such as Hebrew or Czech, created through AI augmentation rather than decades of manual effort.

These viewing patterns reveal that translated articles represent high-value additions to Wikipedia's knowledge base. In Wikipedia's characteristic long-tail distribution, where the median article receives only 10–20 monthly views, translated articles consistently achieve 80–90 views, placing them in the 75th–80th percentile of engagement. This indicates that AI-facilitated content addresses genuine knowledge gaps rather than creating low-quality filler.

For smaller language communities, the acceleration is transformative. Consider Bengali Wikipedia, serving almost 300 million speakers but having accumulated only 80,000 articles by 2019. The AI-enabled acceleration compresses what would require decades of manual effort into just two years, fundamentally altering the trajectory of knowledge development for underserved language communities.

The significance extends beyond immediate metrics through positive externalities. Expanded Wikipedia access influences scientific research (Thompson and Hanley 2018), judicial decisions (Thompson et al. 2024), and economic outcomes (Hinnosaar et al. 2023). These multiplier effects amplify the social value of expanded knowledge access beyond the direct readership numbers.

In summary, these aggregate results provide compelling evidence for AI's democratizing potential. The deployment of advanced machine translation drove massive productivity gains while maintaining quality

and significantly expanding global knowledge access. However, this aggregate view masks critical disparities in how these benefits are distributed, which we examine next.

7. Distributional Benefits: AI Democratization Paradox

The previous section established a significant aggregate increase in AI-assisted content creation, supporting the technological optimism view. However, aggregate effects can mask critical disparities. We now examine the distributional impact of this AI deployment, investigating the tension between democratizing forces and structural reproduction.

7.1. Heterogeneous Impacts Across Languages

We next examine how the benefits of AI deployment are distributed across Wikipedia's diverse language landscape. This section tests the *structural reproduction* hypothesis: even when communities gain equal technical access to better translation technology, do those with greater complementary resources capture disproportionate gains?

7.1.1. Language-Level Treatment Effects. Our counterfactual imputation approach produces an estimated post-integration effect for each treated language edition. Figure 11 plots these language-level effects, sorted from largest to smallest.

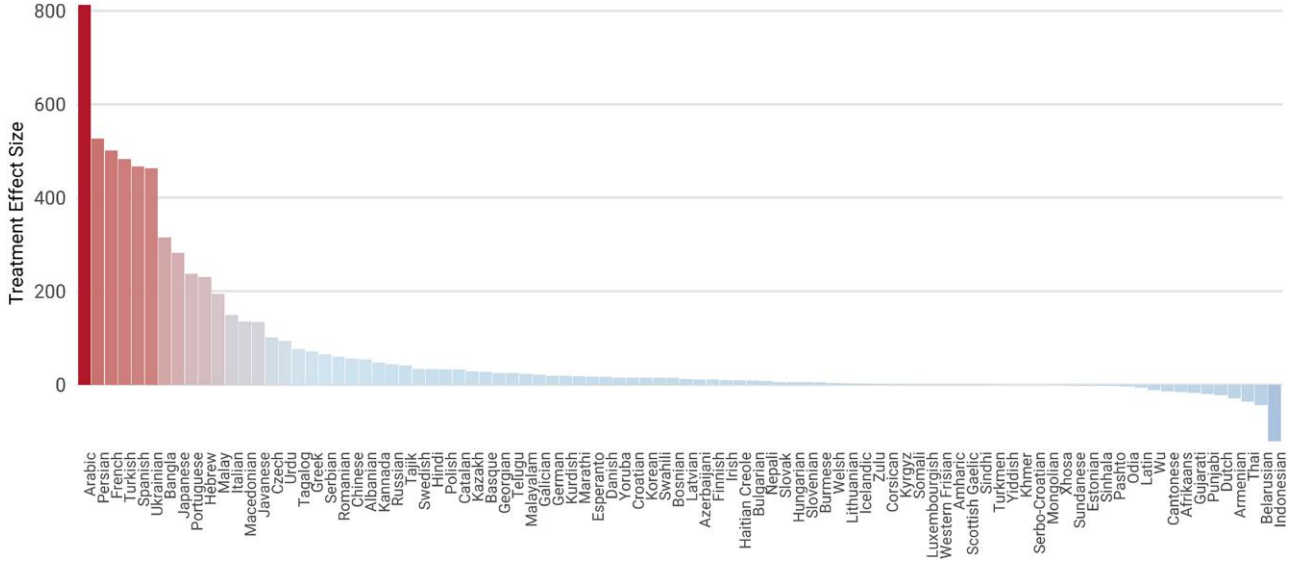
The central result is the *shape of the distribution*. Estimated effects are highly right skewed: a small number of "superwinner" languages realize very large gains, whereas the majority of language editions experience only modest changes. This pronounced skew is the empirical signature of the concentration mechanism we theorize—although AI-enabled translation increases output on average, the resulting benefits accrue disproportionately to better-resourced language communities rather than diffusing evenly across the ecosystem. This pattern motivates a systematic examination, undertaken next, of which language-level characteristics are associated with larger gains.

Because the figure is meant to summarize the distribution (not to support inference about any single language), individual point estimates, including occasional negative values, should be interpreted cautiously.[8]

7.1.2. Unequal Returns to AI Deployment. To probe the determinants of heterogeneity, we analyze treatment effects along three dimensions of community resources: (i) editorial capacity (active editors), (ii) readership networks (pageviews), and (iii) knowledge base size (existing articles). We partition language editions into 10 deciles for each dimension and estimate a heterogeneous difference-in-differences model that interacts treatment with decile indicators.

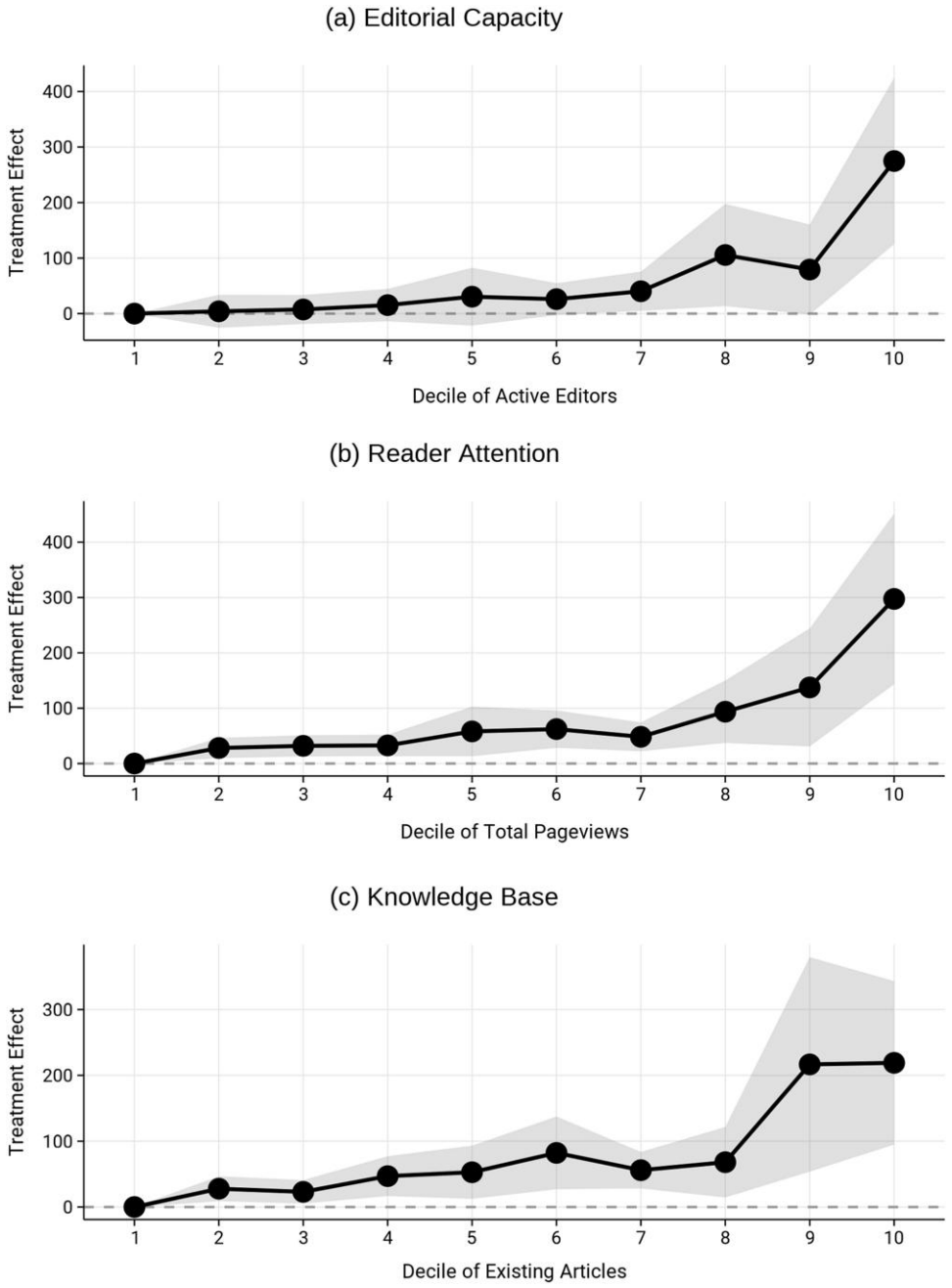
Figure 12 reveals a pronounced gradient. Whereas most communities benefit from Google Translate, well-resourced language editions gain far more. Languages in the top decile realize effects roughly three to four times those of mid-tier communities and more than an order of magnitude larger than the smallest editions. The pattern is consistent across all three resource dimensions. Whereas confidence intervals

Figure 11. (Color online) Language-Level Treatment Effects from AI-Enabled Translation



Notes. Each bar reports the estimated treatment effect (post-integration outcome minus the imputed counterfactual) for one of the 90 treated language editions, sorted from largest to smallest. Values reflect changes in translation-mediated article creation during the post-period; positive bars indicate increases relative to the counterfactual without the tool.

Figure 12. Concentration of AI-Induced Gains Across Language Communities



Notes. Heterogeneous difference-in-differences estimates by deciles of complementary resources: knowledge base size (existing articles), editorial capacity (active editors), and readership networks (pageviews). Points/lines show estimated effects relative to the pre-integration baseline; shaded bands denote 95% confidence intervals. Top-decile languages experience gains three to four times larger than mid-tier communities and over 10× larger than the smallest editions.

for adjacent deciles may overlap, formal joint tests reject the null hypothesis of equal effects ($F = 2.73$, $p = 0.003$), and pairwise comparisons confirm that top and bottom deciles are statistically distinguishable (difference = 275 articles, $t = 3.60$, $p < 0.001$). The monotonic gradient—with effects increasing consistently from bottom to top deciles—provides further support for systemic variation.

These findings support the structural reproduction hypothesis and document a "rich get richer" phenomenon in AI deployment. Technological benefits systematically concentrate among already advantaged communities—large editions such as French, Spanish, Italian, and Arabic that were already central hubs of translation activity prior to the intervention

(see Figures 4 and 5). Although AI lowers language-related barriers for all, communities with greater complementary resources are better positioned to translate, review, and integrate content at scale.

We interpret these findings as robust conditional correlations demonstrating systematic variation in community responses to AI deployment, rather than strict causal mechanisms of community size per se. The key insight is that technological capabilities interact with existing social structures to produce differential outcomes.

Moreover, two sets of additional analyses in the Online Appendix probe the robustness of these findings. First, one might wonder whether the gradient reflects differences in Google Translate quality across languages rather than community resources. Section E.3 shows that treatment effects do not vary systematically with translation quality—the gradient is essentially flat across quality deciles, ruling out this alternative explanation. Second, our main results document absolute effects; a more stringent test asks whether well-resourced communities also gain more *relative to their size*. Online Appendix D presents analyses scaling outcomes by resource base and examining proportional gains. The gradient persists, confirming that the rich-get-richer phenomenon operates on both absolute and relative dimensions.

7.1.3. Centralized Sources, Diversified Access. The preceding analysis reveals that well-resourced communities capture disproportionate gains from AI deployment. But how do these heterogeneous effects reshape the broader structure of cross-linguistic knowledge flows?

Figure 13 reveals two simultaneous yet seemingly contradictory trends following Google Translate integration. Figure 13(a) documents increasing centralization of sources: English Wikipedia's share as the origin of translations rises from 68% to 81%, intensifying its dominance as the world's primary knowledge exporter.

Figure 13(b), however, shows concurrent democratization of access. The diversity of target languages—measured by Shannon entropy—rises markedly after integration, indicating that translated knowledge now reaches a broader array of recipient languages than before. Smaller language communities that previously received little translated content are now being served.

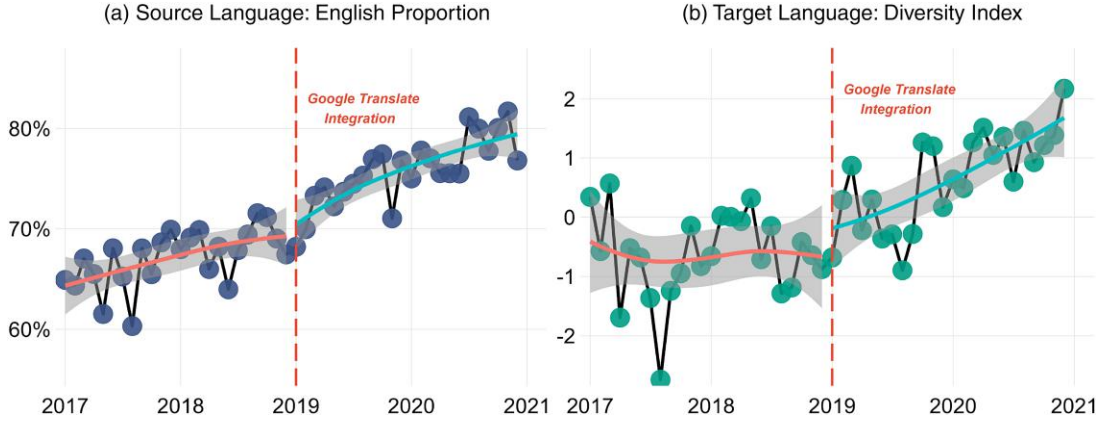
This pattern embodies the AI democratization paradox at the system level. AI produces a more centralized *origin* of knowledge flows—consolidating English Wikipedia's role as the dominant source—while simultaneously broadening *access* across recipient languages. The same technology that amplifies concentration on the supply side enables democratization on the demand side. Whether one views this as progress depends on which dimension one prioritizes: the diversification of who can access knowledge or the homogenization of where that knowledge originates.

7.2. Content Representation

Beyond volume effects, we now examine *what* content gets translated, investigating whether efficiency gains address systemic representation gaps (democratization through agency) or merely reproduce existing biases (structural reproduction).

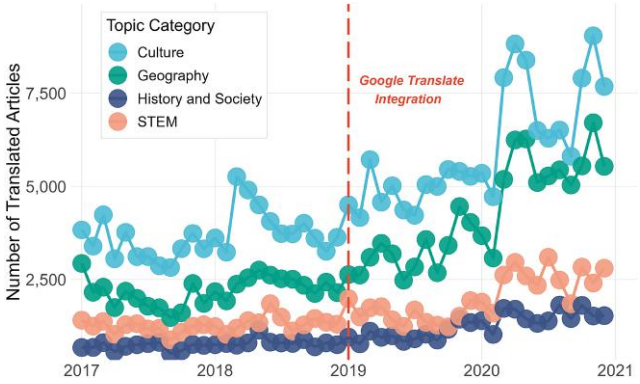
Figure 14 shows heterogeneous growth across knowledge domains. Cultural and geographical articles show the most pronounced increases and collectively represent 75% of translations. Given that these are also critical domains for representation, we focus our subsequent analysis on these areas.

Figure 13. (Color online) The Paradox of Global Knowledge Flows: More Centralized Sources, More Diversified Access



Notes. Evolution around the Google Translate integration (vertical red line). (a) Source centralization: English Wikipedia's share as the origin of translation rises from 68% to 81%. (b) Access diversification: the Shannon entropy of target languages increases, indicating that translated knowledge reaches a broader array of recipient languages.

Figure 14. (Color online) Heterogeneous Growth Across Knowledge Domains, Dominated by Cultural and Geographical Content



Notes. Monthly translation volume across major content categories before and after Google Translate integration (red line). The growth is heterogeneous across domains, with Cultural and Geographical articles showing the most pronounced increases. These two categories collectively represent 75% of all translations post-integration.

7.2.1. Detecting Human Agency via Simulation Benchmarks. To rigorously assess agency, we employ a simulation benchmark analysis. We establish a counterfactual baseline by simulating the translation patterns that would emerge if editors selected articles according to the distribution of the existing resource pool. This null model represents the expected outcome if translation patterns simply reproduced existing structural imbalances without intentional intervention. By comparing the actual observed translation patterns (estimated via DiD) against this simulated benchmark, we can detect active efforts to address representation gaps.

7.2.2. Gender Representation: Democratization with Limitation. We first examine gender representation in biographical articles, where women and nonbinary individuals face well-documented underrepresentation. English Wikipedia's baseline distribution starkly illustrates this imbalance: 82.2% male, 17.7% female, and 0.1% nonbinary biographies.

If editors passively reproduced these patterns, translated content would mirror these proportions. Instead, actual post-integration translations show substantial rebalancing: 62.4% male, 37.2% female, and 0.4% nonbinary. Figure 15 compares our DiD estimates against simulated benchmarks based on random selection from the existing source pool.[9]

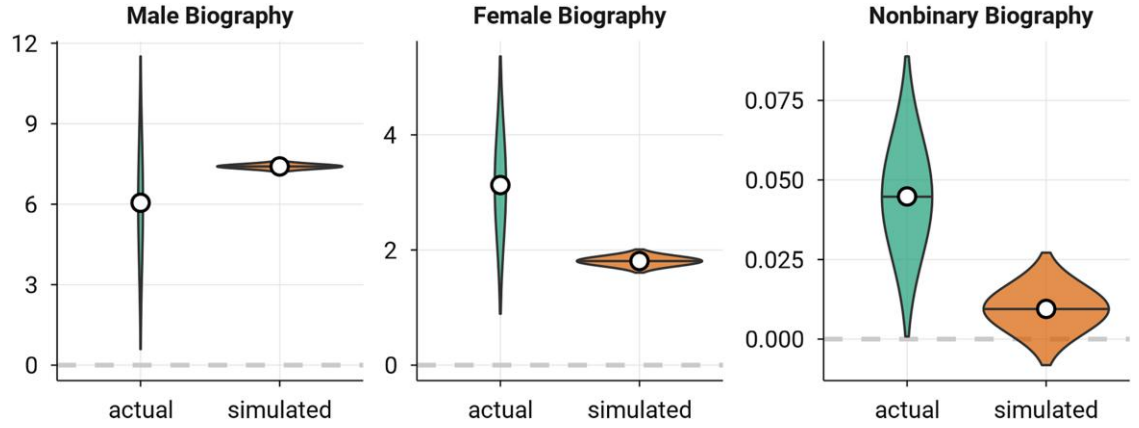
The results reveal significant deviations from the null model. Female biography translations occur at approximately twice the expected rate, whereas nonbinary translations also substantially exceed their benchmark. Male biographies fall significantly below simulated expectations.

These patterns provide suggestive evidence of human agency in leveraging AI tools. Editors are actively using the efficiency gains from advanced machine translation to prioritize underrepresented content. However, the persistence of male biographical dominance in absolute terms suggests that technological solutions alone, although beneficial, may require complementary interventions to fully address deeply embedded structural inequalities.

7.2.3. Geographic Representation: Structural Constraints. Applying the same methodology to geographical content across eight global regions (defined by Wikimedia Foundation; see Section E.7 in the Online Appendix) reveals a more complex picture. Figure 16 shows mixed evidence when comparing actual and simulated translation patterns across regions.

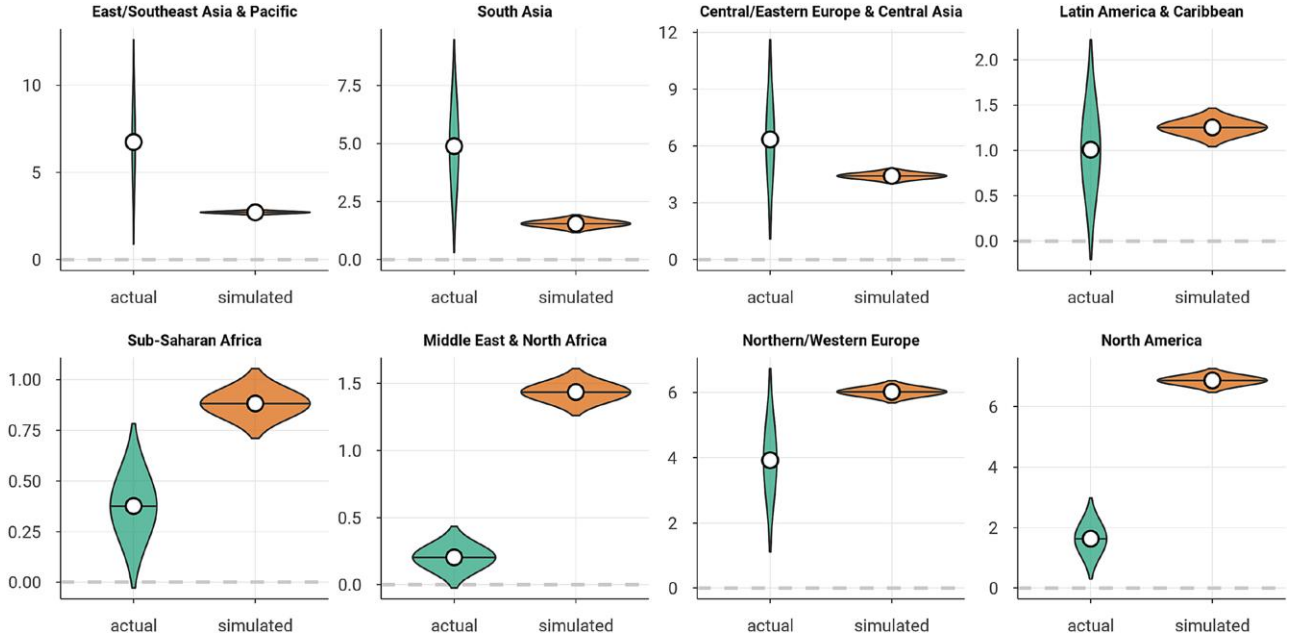
Some progress emerges: East/Southeast Asia and Central/Eastern Europe significantly exceed simulated

Figure 15. (Color online) Evidence of Human Agency: Leveraging AI to Address Gender Representation Gaps



Notes. Comparison of observed post-integration translation patterns (DiD estimates with 95% CIs) against simulated benchmarks (violin plots representing 500 replications). Benchmarks represent expected outcomes based on random selection from the existing, biased source pool (e.g., English Wikipedia: 17.7% female). Female biographies are translated at approximately twice the expected rate (actual observed rate: 37.2%), and nonbinary translations also significantly exceed their benchmarks. Male biographies fall significantly below expectations. This suggests editors actively used AI efficiency gains to prioritize underrepresented content.

Figure 16. (Color online) The Limits of Agency: Structural Constraints Hinder Progress in Geographic Representation



Notes. Comparison of observed translation patterns (DiD estimates) against simulated benchmarks for geographical content across global regions. Whereas some rebalancing occurs (East/Southeast Asia and Central/Eastern Europe exceed benchmarks; North America and Western Europe fall below), regions with the greatest need (e.g., sub-Saharan Africa, Middle East, and North Africa) show minimal absolute gains and do not significantly deviate from random selection. This highlights how structural constraints (lack of source material and local editors) limit the impact of AI.

baselines, whereas North America and Western Europe fall below expectations, suggesting intentional rebalancing. However, regions with the greatest needs—sub-Saharan Africa, the Middle East, and North Africa—show minimal absolute gains (0.4 and 0.3 articles/month) that don't significantly deviate from random selection.

This reveals the boundaries of human agency. Where source content exists and some editorial capacity remains, editors can leverage AI for bias correction. But in regions lacking both source material and local editors, even advanced AI combined with good intentions cannot overcome structural constraints.

7.3. Summary of the Paradox

Our distributional and content analyses reveal what we term the *AI democratization paradox*: the simultaneous operation of democratizing and concentrating forces following AI deployment. This paradox manifests across three interconnected key dimensions:

7.3.1. The Mechanism-Outcome Contradiction. Google Translate operates through an inherently democratizing mechanism—expertise unbundling that universally lowers language barriers. Yet this barrier-reducing technology produces concentrating outcomes, with well-resourced communities capturing gains at least three to four times larger than their disadvantaged counterparts. The paradox lies in how a universally accessible tool amplifies, rather than alleviates, structural inequalities.

7.3.2. Coexistence of Opposing Forces. The same technological intervention generates contradictory dynamics simultaneously. Democratizing forces are evident: knowledge access in more diverse languages, and female biographies are translated at twice the expected rate. Yet concentrating forces operate in parallel: English Wikipedia's dominance increases from 68% to 81%, and regions with the highest need see minimal benefit. AI functions simultaneously as both an equalizer and amplifier of inequality.

7.3.3. The Equal Access Fallacy. Perhaps most striking, our findings challenge the intuitive belief that equal access to democratizing technology should narrow gaps between communities. Instead, identical access produces sharply divergent outcomes. This reveals a deeper truth: the binding constraint on democratization isn't technological capability but complementary resources—editorial capacity, existing knowledge, organizational infrastructure—that remain fundamentally unequal.

This paradox demonstrates that AI's distributional impact emerges not from its technical features but from the complex interaction between technological affordances and social structures. Whereas AI can lower barriers and enable human agency to address

inequalities, these democratizing effects are systematically mediated—and often overwhelmed—by the unequal distribution of complementary resources. Understanding this dynamic is essential for designing AI deployments that achieve, rather than undermine, equitable outcomes.

8. Conclusion and Implications

This study provides the first large-scale empirical evidence on how AI deployment reshapes knowledge production in decentralized platforms. By leveraging Wikipedia's integration of Google Translate as a natural experiment across more than 100 diverse language communities, we examined the impact of advanced AI on cross-linguistic content creation. Our findings reveal a central tension in AI deployment within collaborative systems, which we term the *AI democratization paradox*.

Our empirical analysis demonstrates that AI deployment simultaneously unleashes both democratizing and concentrating forces. The democratizing forces are substantial: we found a more than double increase in aggregate translation-mediated content creation. This generates an estimated 12.3 million additional pageviews over two years, critically distributed across a diverse array of target languages, making knowledge accessible to speakers of previously underserved linguistic communities worldwide. Moreover, we observed clear evidence of human agency leveraging AI to address systemic inequalities, with female biographies translated at twice the expected rate.

However, these gains are profoundly mediated by strong concentrating forces. Benefits disproportionately flow to well-resourced language communities with greater editorial capacity, larger existing knowledge bases, and established reader networks. These communities experienced gains three to four times larger than smaller editions, whereas regions with the highest need, such as sub-Saharan Africa, saw minimal impact. This paradox highlights that AI's impact critically depends on the interaction between technological capabilities and existing social structures.

Our study advances understanding of AI adoption, digital platforms, and inequality by moving beyond "AI increases productivity" narratives to demonstrate how socio-technical systems shape AI's distributional impact. We provide a conceptual framework reconciling technological optimism and structural reproduction views, showing that both mechanisms operate simultaneously. We empirically validate that technological capabilities alone cannot overcome structural inequalities, as the capacity to leverage AI depends on existing human capital and organizational infrastructure. Finally, we highlight human agency's role in steering AI toward equitable outcomes while identifying constraints on that agency.

Our findings offer actionable insights for organizations deploying AI in collaborative systems. Providing equal access to AI tools is necessary but insufficient for equitable outcomes. Platforms must recognize that "one-size-fits-all" AI deployment risks amplifying existing disparities. To realize AI's democratizing potential, technological solutions must be coupled with complementary interventions supporting under-resourced communities through targeted training, capacity building, and infrastructure support. Organizations should evaluate AI impact through a distributional lens rather than focusing solely on aggregate productivity metrics. Policymakers addressing the digital divide must recognize that equitable access to information requires fostering both technological access and the human and organizational capacity to leverage that technology effectively.

8.1. Avenues for Future Research

Our findings open several promising directions that extend both theoretical understanding and practical applications of AI in collaborative systems. First, could platform operators mitigate AI's concentrating effects through targeted interventions? Future research should examine how governance choices, such as providing editorial support, capacity-building programs, or differential incentives to under-resourced communities, can shape AI's distributional impact. This connects our framework to the growing literature on platform governance, offering actionable insights for organizations deploying AI in decentralized systems.

Second, as AI systems become increasingly autonomous, will the democratization paradox intensify or evolve? Advanced generative models might prove more democratizing by reducing dependence on complementary human resources, or they may create new inequality mechanisms through AI skill gaps and sophisticated content verification challenges. Understanding how our paradox transforms with increasingly capable AI represents a promising research frontier.

Third, does AI-assisted content creation, which predominantly flows from dominant languages such as English, promote cultural homogenization or perpetuate embedded biases? Future work should examine whether efficiency gains from AI come at the cost of cultural authenticity and diverse knowledge traditions. Exploring these nuanced dimensions of value in global knowledge commons offers rich opportunities for interdisciplinary collaboration.

Fourth, future research should complement our supply-side analysis by structurally estimating demand. By leveraging exogenous variation in content availability to model latent demand curves, scholars could quantify the aggregate welfare effects of AI translation and rigorously examine how this consumer surplus is distributed across diverse language communities.

In conclusion, realizing AI's democratizing potential requires sustained investigation of the complex interplay between technological capabilities, social structures, and governance mechanisms. Recognizing and addressing this challenge—by coupling technological innovation with targeted structural support—is essential for ensuring that the AI revolution contributes to a more equitable digital future.

Endnotes

[1] See the official announcement at https://wikimediafoundation.org/news/2019/01/09/you-can-now-use-google-translate-to-translate-articles-on-wikipedia/.

[2] These are defined as monthly volumes exceeding 20 times a language's historical average.

[3] We verify robustness across multiple imputation methods provided by the "fect" R package, including fixed effects counterfactual, interactive fixed effects, and matrix completion approaches, all yielding consistent results.

[4] We calculate the monthly deletion ratio by aggregating across all languages: (number of articles translated in month M that were eventually deleted)/(total number of articles translated in month M). Aggregating results across all languages provides a stable monthly metric, as smaller language editions display high variability because of low monthly translation volumes.

[5] Here, we examine substitution versus expansion at the individual article level. For an analysis at the macro–language level, see Section F of the Online Appendix.

[6] Whereas our post-translation readership evidence suggests that demand exists for translated content, we cannot estimate the full distribution of unmet demand across language pairs, nor can we claim that translating any arbitrary article would generate equivalent readership.

[7] Calculation: 71.2 articles/month/language × 90 treated languages × 24 months × 80 views/month = 12.3 million views.

[8] Language-level estimates inherit uncertainty from both realized post-period outcomes and the imputed counterfactual. Our implementation reports uncertainty for aggregate and binned analyses (e.g., event studies and decile-by-resource interactions), but it does not yield reliable standard errors/confidence intervals for each single-language estimate. Accordingly, a small number of negative point estimates are expected from estimation noise.

[9] See Table A.6 in Section E.5 of the Online Appendix for detailed regression results.

References

Aaltonen A, Seiler S (2016) Cumulative growth in user-generated content production: Evidence from Wikipedia. *Management Sci.* 62(7):2054–2069.

Agrawal A, Gans J, Goldfarb A (2019) *The Economics of Artificial Intelligence: An Agenda* (University of Chicago Press, Chicago).

Agrawal A, Gans J, Goldfarb A (2022) *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence* (Harvard Business Review Press, Boston).

Athey S, Bayati M, Doudchenko N, Imbens G, Khosravi K (2021) Matrix completion methods for causal panel data models. *J. Amer. Statist. Assoc.* 116(536):1716–1730.

Benkler Y (2002) Coase's penguin, or, Linux and "the nature of the firm". *Yale Law J.* 112(3):369–446.

Brynjolfsson E, Hui X, Liu M (2019) Does machine translation affect international trade? Evidence from a large digital platform. *Management Sci.* 65(12):5449–5460.

Brynjolfsson E, Li D, Raymond L (2025) Generative AI at work. *Quart. J. Econom.* 140(2):889–942.

Callaway B, Sant'Anna PH (2021) Difference-in-differences with multiple time periods. *J. Econom.* 225(2):200–230.

Capraro V, Lentsch A, Acemoglu D, Akgun S, Akhmedova A, Bilancini E, Bonnefon J-F, et al. (2024) The impact of generative artificial intelligence on socioeconomic inequalities and policy making. *PNAS Nexus* 3(6):pgae191.

Cazzaniga M, Jaumotte MF, Li L, Melina MG, Panton AJ, Pizzinelli C, Rockall EJ, Tavares MMM (2024) *Gen-AI: Artificial Intelligence and the Future of Work* (International Monetary Fund, Washington, DC).

Cengiz D, Dube A, Lindner A, Zipperer B (2019) The effect of minimum wages on low-wage jobs. *Quart. J. Econom.* 134(3):1405–1454.

Cohen WM, Levinthal DA (1990) Absorptive capacity: A new perspective on learning and innovation. *Admin. Sci. Quart.* 35(1): 128–152.

De Chaisemartin C, d'Haultfoeuille X (2020) Two-way fixed effects estimators with heterogeneous treatment effects. *Amer. Econom. Rev.* 110(9):2964–2996.

Faraj S, von Krogh G, Monteiro E, Lakhani KR (2016) Special section introduction—Online community as space for knowledge flows. *Inform. Systems Res.* 27(4):668–684.

Forman C, Goldfarb A, Greenstein S (2012) The internet and local wages: A puzzle. *Amer. Econom. Rev.* 102(1):556–575.

Goldfarb A, Tucker C (2019) Digital economics. *J. Econom. Literature* 57(1):3–43.

Goodman-Bacon A (2021) Difference-in-differences with variation in treatment timing. *J. Econom.* 225(2):254–277.

Graham M, Hogan B, Straumann RK, Medhat A (2014) Uneven geographies of user-generated information: Patterns of increasing informational poverty. *Ann. Assoc. Amer. Geographers* 104(4):746–764.

Hinnosaar M, Hinnosaar T, Kummer M, Slivko O (2023) Wikipedia matters. *J. Econom. Management Strategy* 32(3):657–669.

Holland PW (1986) Statistics and causal inference. *J. Amer. Statist. Assoc.* 81(396):945–960.

Jensen R (2007) The digital provide: Information (technology), market performance, and welfare in the South Indian fisheries sector. *Quart. J. Econom.* 122(3):879–924.

Johnson I, Gerlach M, Sáez-Trumper D (2021) Language-agnostic topic classification for Wikipedia. *WWW'21: Companion Proc. Web Conf. 2021* (Association for Computing Machinery, New York), 594–601.

Miquel-Ribé M, Laniado D (2020) The Wikipedia Diversity Observatory: A project to identify and bridge content gaps in Wikipedia. *OpenSym'20: Proc. 16th Internat. Sympos. Open Collaboration* (Association for Computing Machinery, New York), 1–4.

Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge University Press, Cambridge, UK).

Parker GG, Van Alstyne MW, Choudary SP (2016) *Platform Revolution: How Networked Markets Are Transforming the Economy and How to Make Them Work for You* (W. W. Norton & Company, New York).

Rochet J-C, Tirole J (2003) Platform competition in two-sided markets. *J. Eur. Econom. Assoc.* 1(4):990–1029.

Scheidel W (2024) Beyond technology and wages: Power and the history of inequality. *Oxford Open Econom.* 3(Supplement_1): i212–i216.

Sun L, Abraham S (2021) Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *J. Econom.* 225(2):175–199.

Teh T-H (2022) Platform governance. *Amer. Econom. J. Microeconom.* 14(3):213–254.

Thompson N, Hanley D (2018) Science is shaped by Wikipedia: Evidence from a randomized control trial. Preprint, submitted September 20, 2017, https://doi.org/10.2139/ssrn.3039505.

Thompson N, Luo X, McKenzie B, Richardson E, Flanagan B (2024) User-generated content shapes judicial reasoning: Evidence from a randomized control trial on Wikipedia. *Inform. Systems Res.* 35(4):1948–1964.

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need. von Luxburg U, Guyon I, Bengio S, Wallach H, Fergus R, eds. *NIPS'17: Proc. 31st Internat. Conf. Neural Inform. Processing Systems* (Curran Associates Inc., Red Hook, NY), 6000–6010.

Von Hippel E (2006) *Democratizing Innovation* (MIT Press, Cambridge, MA).

Wagner C, Graells-Garrido E, Garcia D, Menczer F (2016) Women through the glass ceiling: Gender asymmetries in Wikipedia. *EPJ Data Sci.* 5(1):5.

Wessels B (2013) The reproduction and reconfiguration of inequality: Differentiation and class, status and power in the dynamics of digital divides. Ragnedda M, Muschert GW, eds. *The Digital Divide* (Routledge, London), 17–28.

Wu Y, Schuster M, Chen Z, Le QV, Norouzi M, Macherey W, Krikun M, et al. (2016) Google's neural machine translation system: Bridging the gap between human and machine translation. Preprint, submitted September 26, 2016, https://arxiv.org/abs/1609.08144.

Zhang X, Zhu F (2011) Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *Amer. Econom. Rev.* 101(4):1601–1615.

Zhu K, Walker D, Muchnik L (2020) Content growth and attention contagion in information networks: Addressing information poverty on Wikipedia. *Inform. Systems Res.* 31(2):491–509.