The Divergent Influence of Generative AI on Knowledge Dynamics Within Stack Exchange

Stella Jia

May 2024

Abstract

Throughout history, humans have employed various methods to gather and disseminate knowledge. From oral traditions and written texts to libraries and the internet, the means of knowledge sharing have continually evolved. With the emergence of generative AI like ChatGPT, how will our online knowledge ecosystems change? I analyze user engagement on Stack Exchange, a platform covering diverse domains, following the introduction of ChatGPT in November 2022. I categorize knowledge types into tacit, explicit, and hybrid, as defined by Polanyi's paradox. While effects vary across domains, there is no significant impact on the level of knowledge types.

1

Contents

1 Introduction

The rise of digital information platforms marked a significant shift in the dissemination of knowledge. What once was confined to human memory and physical records now exists on things beyond our own minds and locations, enabling open and collaborative knowledge exchange on an unprecedented global scale. Platforms like Wikipedia, Reddit, and Stack Exchange became vital repositories of information, fueled by contributions from individuals worldwide.

While digital information platforms have served as vital repositories of knowledge, the emergence of generative AI models like ChatGPT represents a new frontier in knowledge interaction and creation. Existing labor economics research has explored ChatGPT's capabilities and limitations across various skills and occupations (Felten et al., 2021). One area where ChatGPT has demonstrated remarkable proficiency is code writing (Chen et al., 2021; Austin et al., 2021), prompting investigations into its potential impact on technical knowledge sharing platforms like Stack Overflow (Burtch et al., 2023).

However, ChatGPT's influence extends beyond the realm of technical knowledge. Studies have highlighted its utility in areas such as writing, market research (Brand et al., 2023), and digital artwork (Zhou and Lee, 2024). Consequently, it becomes imperative to examine how this technology might reshape the dynamics of knowledge dissemination across diverse domains, extending beyond the technical sphere. In this vein, the work of Marshall McLuhan, "The Medium is the Message," presents a compelling perspective on how the medium through which ideas are conveyed can profoundly impact their interpretation and assimilation (McLuhan, 1994). McLuhan posited that the form of a medium embeds itself in the message, creating "massive extensions of the human mind and body" (p. 26). How will the medium of generative AI shape the dissemination and understanding of knowledge across different domains on our existing digital information platforms?

While open knowledge sharing platforms like Wikipedia and Stack Overflow are fueled by decentralized, crowdsourced contributions from a multitude of individuals, ChatGPT

represents a centralized source of information generated by a single, opaque AI model. This fundamental difference in the mediums through which knowledge is conveyed could have implications for how that knowledge is perceived, interpreted, and assimilated by users.

To contribute to this discourse, I leverage publicly available data from the Stack Exchange network, to analyze the changes in question volumes across different topic areas before and after the introduction of ChatGPT in November 2022. I hope to shed light on the immediate and measurable effects of generative AI on user interaction within these knowledge ecosystems across different domains. Domains exhibiting decreased question volumes may indicate that users are finding satisfactory answers from ChatGPT itself, while increased volumes could suggest that the AI's outputs are stimulating further curiosity and prompting users to seek more in-depth or nuanced knowledge from the community. Additionally, variations in volume changes across domains may reveal fundamental differences in how generative AI interacts with different types of knowledge, such as factual information (coding, statistics, physics, etc) versus subjective or interpretive topics (philosophy, religion, politics, etc).

While this study may not directly address the profound philosophical and societal implications of generative AI's impact on knowledge creation and dissemination, it serves as a stepping stone. By empirically quantifying the immediate effects on user behavior and engagement across different types of knowledge, this research lays the groundwork for further explorations into the broader impact and opportunities presented by this technology.

2 Literature Review

2.1 Generative AI and Digital Platforms

The emergence of generative AI has sparked a wide discussion on its potential to transform various aspects of society, from the labor market (Felten et al., 2021) to the dynamics of digital platforms. There is a growing body of research investigating the impact of generative AI on digital information platforms. Such digital platforms ranging from open knowledge

repositories like Wikipedia, Reddit, and Stack Overflow, to monetized content providers like the New York Times[1] and Getty Images[2] —serve as key training data to AI models which require more detailed investigation to establish appropriate policies (Gans, 2024).

Among open knowledge-sharing platforms, del Rio-Chanona et al. (2023) reveals a 16% decrease in weekly posts on Stack Overflow after the release of ChatGPT. The decrease is especially evident in popular programming languages like Python and Javascript, however, for more niche languages like CUDA, which ChatGPT may not have as much training data on, we actually see an increase in posts after ChatGPT.

Burtch et al. (2023) provide evidence that straightforward coding questions, characterized by their concrete nature, are readily replaceable due to their widespread availability on platforms such as GitHub, which are utilized by models like ChatGPT for training purposes. Conversely, topics on Stack Overflow that are less affected tend to be highly contextual, where solutions cannot simply be copied and pasted. These specific contexts often do not appear in the training datasets, making automated solutions less feasible. Additionally, the study highlights a notable decline in the quality of answers on Stack Overflow following the introduction of ChatGPT, evidenced by a significant two percentage point increase in the likelihood of answers being downvoted. This decline is attributed primarily to a reduction in the number of questions answered by experts, rather than a decrease in the quality of responses by individual users.

2.2   Theory of Knowledge

To more deeply understand the dynamics between generative AI and knowledge, I looked into defining concepts in the theory of knowledge. Tacit versus explicit knowledge represents a fundamental dichotomy in the field of knowledge management, profoundly influencing how information is communicated, understood, and used in various domains. Tacit knowledge, as

---

[1]New York Times, "New York Times, OpenAI, Microsoft Lawsuit," December 27, 2023, `https://www.nytimes.com/2023/12/27/business/media/new-york-times-open-ai-microsoft-lawsuit.html`.

[2]BakerHostetler, "Getty Images v. Stability AI," accessed May 5, 2024, `https://www.bakerlaw.com/getty-images-v-stability-ai/`.

initially conceptualized by Michael Polanyi is inherently personal, deeply rooted in individual experiences, and often difficult to articulate. It encompasses skills, ideas, and intuitions that are acquired through personal involvement and context-specific interactions, making it challenging to formalize or transmit through conventional means (Polanyi, 1966). In contrast, explicit knowledge is easily codifiable, transmittable through systematic language, and can be readily documented, stored, and accessed (Nonaka and Takeuchi, 1995). This type of knowledge is often found in manuals, databases, and scientific formulas, where its clarity and universality allow for straightforward communication and application.

In the context of online knowledge-sharing platforms, understanding how tacit and explicit knowledge are shared and transformed can offer insights into the effectiveness of these platforms in facilitating meaningful exchanges versus straightforward information dissemination. This distinction also bears implications for the design of AI-driven tools like Chat-GPT, which might excel in processing and generating explicit knowledge but face challenges in capturing the nuanced, context-dependent subtleties of tacit knowledge. Thus, the tacit vs. explicit knowledge framework not only enriches our understanding of knowledge dynamics but also guides the development of technologies and strategies for effective knowledge management in the digital age.

In considering the distinction between tacit and explicit knowledge types, Burtch et al. (2023) found that context-dependent programming topics such as Azure, AWS, Excel, and Docker experienced less significant declines in user engagement post the introduction of AI tools like ChatGPT, compared to more fundamental programming languages like Python, JavaScript, and ReactJS. Such technical topics like Python are representative of explicit knowledge while more rare and context-dependent ones like AWS are similar to tacit knowledge. This pattern highlights the sensitivity of explicit knowledge, which is highly dependent on personal context, underscoring the challenges AI faces in effectively replicating the nuanced transfer of such knowledge. Burtch et al. (2023) findings intertwine with Nonaka and Takeuchi (1995) theory on the difficulty of converting tacit knowledge, suggesting a limited

impact of AI on context-rich discussions.

2.3 Research Gaps

Existing research offers valuable insights, yet gaps remain in our understanding of the nuanced ways in which generative AI is transforming information-seeking behaviors across different knowledge types. Initial observations suggest a decline in platform engagement and varying effects across different programming languages and software tools, indicating broader trends that warrant further investigation. Specifically, it is crucial to explore the impact of large language models (LLMs) on different types of knowledge discussed, particularly outside of technical fields. Such an analysis could provide a more comprehensive view of generative AI's influence on the knowledge economy, aiding digital platforms in devising strategies to navigate the burgeoning era of generative AI. As these AI models evolve, the demand for data to train and enhance their quality increases. However, with much of the accessible data already utilized, the question arises: what sources will fuel the training of future iterations? This paper aims to highlight the value of information platforms by examining how ChatGPT affects various types of knowledge.

3 Data

Stack Exchange is a network of 173 Q&A communities, each dedicated to a specific topic or field. Among these, Stack Overflow stands out as the most well-known, with over 58 million posts primarily focused on technical questions and answers. The Internet Archive stores data from all 173 communities within Stack Exchange from its inception until April 2024, capturing a wide array of interactions and content types ranging from Judaism to Computer Science. There are also exists "Meta" threads which involve discussions around community as a whole and how the software works like Meta StackExchange and Meta StackOverflow.

For the purposes of this study, I have selected all threads that contain over 100 megabytes

Table 1: Total Question Count by Thread from January 18, 2021 to April, 1 2024

Thread	Count	Thread	Count	Thread	Count
Academia	9,861	Electronics	51,485	ServerFault	32,209
Android	7,308	English	13,298	SharePoint	7,917
Apple	19,936	English Language Learners	25,593	Software Engineering	5,425
AskUbuntu	63,862	Gaming	7,774	StackOverflow	3,972,574
Blender	41,418	Hermeneutics	5,900	Stats	49,804
Chemistry	8,480	Judaism	7,633	TeX	50,898
Christianity	4,097	Magento	12,095	User Experience	3,017
CodeGolf	2,165	Meta StackExchange	6,150	WordPress	11,970
CodeReview	9,424	Meta StackOverflow	5,943	Physics	61,473
Computer Science	10,328	Money	7,248	Politics	5,222
DBA	17,994	Philosophy	6,474	Salesforce	26,951
DIY	26,191	Drupal	5,124	SciFi	13,087

of data. This criterion was chosen to ensure that the selected threads have a high level of activity and interaction, thereby providing a representative sample of the community discussions. The thread with the lowest number of questions in the specified timeframe is Drupal with 5,124 questions posted and the highest number of questions is Stack Overflow with over three million questions followed by AskUbuntu with 63,862.

The onset of the COVID-19 pandemic in early 2020 led to unprecedented disruptions across various sectors, including education, which shifted predominantly online. This transition resulted in significant increases in online activity, as evidenced by spikes in user engagement across educational forums like Stack Exchange[3] . Recognizing that these spikes could skew normal interaction patterns, our study focuses on the period after these initial fluctuations. By early 2021 education begins to reach a stabilization (Fabriz et al., 2021) thus I have chosen January 18, 2021, as the starting point. This cutoff is intended to exclude the initial, most volatile months of the pandemic, thereby focusing our analysis on a more normalized period of online interactions within the Stack Exchange communities.

In analyzing the data, I conducted a comparative time series analysis between Stack Overflow and a group of other discussion threads. I charted the number of questions posted on Stack Overflow on one side, against the aggregate count from the other threads on the other side. From 2021 to 2022, Stack Overflow showed a noticeable decline, dropping by

---

[3]Stack Overflow, "How the pandemic changed traffic trends on Stack Exchange sites," Stack Overflow Blog, April 20, 2020, `https://stackoverflow.blog/2020/04/20/pandemic-changed-traffic-trends-stack-exchange-sites/`.
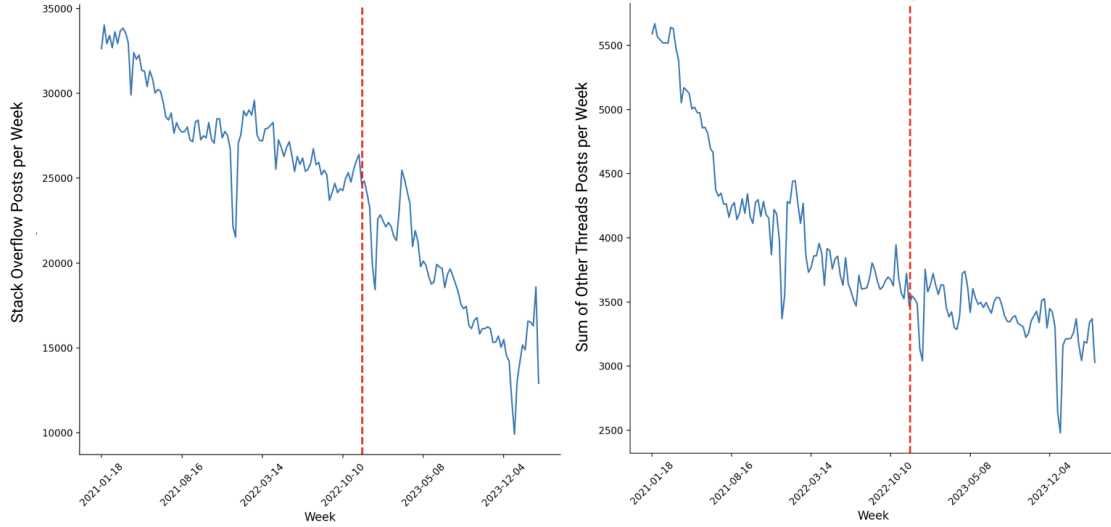
Figure 1: Time Series of Question Volume on Stack Overflow (left) and Other Threads (right)

10,000 posts. In contrast, the other threads collectively decreased by 2,000 posts during the same period. This reduction appears more significant in relative terms due to the smaller total volume of posts on these threads compared to Stack Overflow. While this trend was evident before the introduction of ChatGPT, the decline in post volume at Stack Overflow became more pronounced after ChatGPT's entry, compared to the aggregate of the other threads.

Prior to running the regression, I categorized the threads into tacit vs explicit knowledge types. To aid in this categorization, I referenced Polanyi (1966) defining markers of tacit vs explicit knowledge based on the principle that explicit knowledge can be easily communicated and documented, while tacit knowledge is personal and harder to formalize or communicate. Additionally, I used GPT4 to refine the categorization with zero-shot classification (Wang et al., 2023) and persona prompts (White et al., 2023).

My approach in classifying threads such as Android, Physics, and Computer-Science under explicit knowledge is driven by the structured nature of these fields, which are predominantly based on formal education and documented methodologies. In contrast, areas like User-Experience, Gaming, and WorldBuilding are categorized as tacit knowledge due to their reliance on subjective interpretation and individual proficiency, which are cultivated

Table 2: Categorization of Topics into Explicit, Tacit, and Hybrid Knowledge

Explicit Knowledge	Tacit Knowledge	Hybrid Knowledge
Chemistry	Christianity	DIY
Computer-Science	Hermeneutics	Blender
Physics	Judaism	Drupal
Stats	WorldBuilding	Gaming
TeX	English-Language-Learners	User-Experience
English	SciFi	ServerFault
StackOverflow		Meta.StackExchange
Academia		Meta.StackOverflow
Android		SharePoint
Apple		WordPress
AskUbuntu		SalesForce
CodeGolf		Magento
CodeReview		Politics
DBA		Philosophy
Electronics		
Magento		
Money		
Software-Engineering		

through practice rather than formal instruction.

The binary classification into explicit and tacit knowledge simplifies the complex nature of how knowledge is constructed and communicated. Additionally, many threads may fall into either group depending on the context. For example, the classification of some subjects, especially third party tools/software, like Blender and Drupal illustrates an overlap between explicit and tacit knowledge. While these tools have comprehensive documentation, proficiency in them often requires experiential learning that goes beyond what is written thus I created a new category for this hybrid knowledge.

4   Methodology

The dependent variable, $QuestionVolume_{it}$ represents the total number of posts in a particular thread tagged i during a specific week t. To analyze these dynamics, I employ a Simple Linear Regression (SLR) model that includes interaction terms and fixed effects. The model

is formulated as follows:

$$QuestionVolume_{it} = Thread_i + Post + (Thread_i \cdot Post) + \tau_t + \mu_{it}$$

In this model, $Thread_i$ is a binary indicator representing the presence of the thread tag i, and $Post$ is a binary indicator denoting the introduction of ChatGPT. The interaction term $Thread_i \cdot Post_t$ captures the combined effect of the thread and after ChatGPT's introduction on the discussion volume. The τ_t term represents fixed effects for each month t, controlling for time-specific variations such as weekly trends or seasonal patterns, which corrects for unobserved heterogeneity that might affect all threads uniformly in any given week. The μ_{it} is the error term, accounting for other unobserved influences on the question volume.

To ensure comparability across multiple threads, the count of questions per week was normalized based on each threads mean and standard deviation pre-ChatGPT. The normalization was then applied across the pre and post-ChatGPT data.

5   Results

For simplicity and clarity of the analysis, the table provided focuses on a subset of threads, including both technical and non-technical domains. A comprehensive examination of all analyzed forums is depicted in Figure 2.

Table 3: Summary of OLS Regression Results focusing on Post Interactions with ChatGPT

Variable	Coeff.	S.E.	t	p
Intercept	0.1763	0.098	1.801	0.072
Significant Interactions:				
Apple	-0.8875	0.200	-4.443	<0.001
CodeReview	-1.2145	0.200	-6.080	<0.001
DIY	1.4518	0.200	7.268	<0.001
Judaism	2.1398	0.200	10.712	<0.001
Philosophy	2.9975	0.200	15.005	<0.001
StackOverflow	-2.8449	0.202	-14.117	<0.001
Insignificant Interactions:				
English	-0.7940	0.200	-3.975	<0.001
Electronics	0.0602	0.200	0.301	0.763
Gaming	-0.6499	0.200	-3.254	0.001
User-Experience	0.0709	0.200	0.355	0.723
Meta.StackExchange	-0.2128	0.200	-1.065	0.287
Model Stats:	R^2=0.438, Adj. R^2=0.431, F=56.95, p<0.001, LL=-8135.1			

The regression results reveal significant disparities in how different Stack Exchange threads have responded to the introduction of ChatGPT. Notably, threads such as CodeReview and StackOverflow exhibited substantial negative coefficients (-1.2145 and -2.8449, respectively), with p-values less than 0.001, indicating a significant decline in user interactions post-ChatGPT introduction. This trend suggests that users may be leveraging ChatGPT for coding-related inquiries, likely due to the AI's capacity to provide immediate and accurate coding solutions, thereby reducing the necessity for community-driven support.

Conversely, forums such as DIY, Judaism, and Philosophy displayed significant positive coefficients (1.4518, 2.1398, and 2.9975, respectively), also with p-values less than 0.001.

These findings suggest an increase in user engagement following the rollout of ChatGPT. It appears that the generative nature of ChatGPT might be stimulating further curiosity and encouraging users to delve deeper, prompting more extensive or nuanced discussions within these communities. The positive impact in these domains may indicate that the AI's responses, while informative, often lead to additional questions or fail to fully satisfy users' requests for comprehensive exploration.

In contrast, domains such as Electronics and User-Experience showed insignificant changes in user interactions (coefficients of 0.0602 and 0.0709, respectively, with p-values of 0.763 and 0.723). These results imply a minimal impact of ChatGPT on user engagement within these forums. The lack of significant change could reflect the AI's limited applicability to the specific needs of these communities or a preference among users for traditional forms of interaction over AI-driven assistance.

While I initially expected to see clear patterns in how ChatGPT impacts different types of knowledge on Stack Exchange forums, the results from the regression analysis depicted in Figure 2 do not indicate a distinct pattern. In Table 4, both explicit and tacit exhibit a significant decrease in engagement while hybrid has a non-significant increase on a 0.05 significance level. This lack of divergent effects across tacit, explicit, and hybrid knowledge categories may be due to several underlying factors, one being that each category may not fit cleanly into just one knowledge category.

Table 4: Effect Size by Knowledge Type

Knowledge	Coefficient	P-value	Lower CI	Upper CI
Explicit	-0.972780	0.002741	-1.364692	-0.580868
Hybrid	0.016457	0.012524	-0.375140	0.408055
Tacit	-0.364099	0.004209	-0.755697	0.027498

Looking at individual threads within each knowledge types, threads that deal with tacit knowledge show a mixed range of effects, from slight negatives to mild positives. For instance, forums such as Judaism and Christianity exhibit positive to no effects, with Judaism in
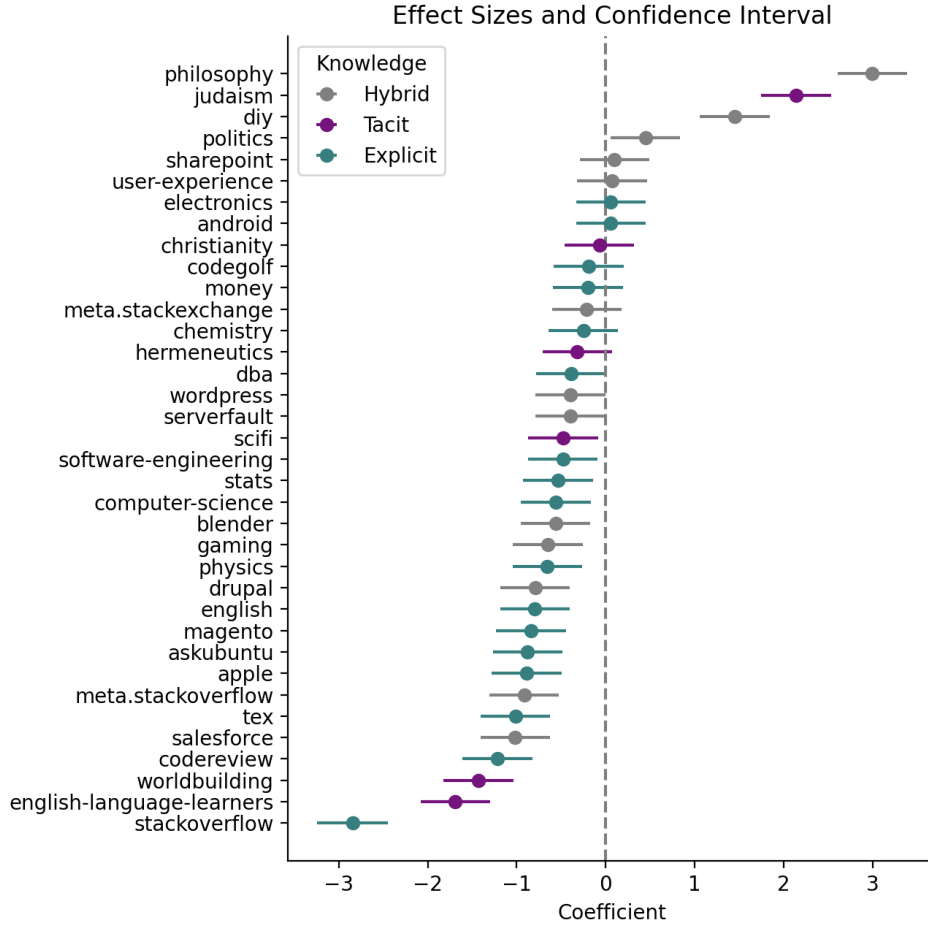
Figure 2: Estimated Change in Question Volume by Thread

particular showing one of the most positive coefficients among tacit knowledge. This suggests that in domains where knowledge is experiential and interpretive, ChatGPT may stimulate further discussion and inquiry, possibly because AI responses lead to more questions than answers.

For threads centered on explicit knowledge, such as StackOverflow, CodeReview and LaTeX, significant negative coefficients were observed. Most explicit threads all have either negative or no effect which matches my initial hypothesis that explicit knowledge areas would show negative impact due to the straightforward nature of the queries that ChatGPT can effectively handle.

The hybrid threads span positive and negative coefficients like Philosophy which is the

most positive and Salesforce which is negative. This variance suggests that the complexity of questions in hybrid forums, which blend factual and experiential queries, might influence how effectively ChatGPT can integrate into and stimulate these community interactions.

6 Discussion

The results of this study suggest a nuanced impact of ChatGPT on various Stack Exchange forums. While the initial hypothesis anticipated clear distinctions in ChatGPT's effects based on the nature of the knowledge, the descriptive results reveal no distinct patterns. This unexpected finding underscores the complexity of AI interactions within knowledge-sharing platforms and suggests that the influence of generative AI cannot be solely attributed to the categorization of knowledge.

In threads dealing with explicit knowledge, such as StackOverflow, the presence of Chat-GPT appears to decrease user interaction, likely due to the AI's ability to provide precise, factual answers, which diminishes the need for further human interaction on a knowledge sharing platform. Conversely, in threads centered around tacit knowledge, such as Philosophy, the data indicates an increase in engagement, possibly because ChatGPT does not suffice and users still benefit from interactions on knowledge sharing platforms. These outcomes highlight the potential of AI to both supplement and supplant human input in online discussions, depending on the context.

6.1 Limitations

One notable limitation of this study is the subjectivity involved in categorizing the forums into different types of knowledge. The classification was based on intuitive assessments rather than a systematic, computational analysis of the text within each forum. This methodological choice could have introduced biases or inaccuracies in categorizing the forums, potentially affecting the analysis outcomes. Future studies could benefit from employing computational

textual analysis to define and categorize forums more rigorously, which would help in achieving more precise and reliable insights.

In examining the impact of ChatGPT on Stack Exchange forums, our study adopts a descriptive approach rather than a causal analysis. This methodology acknowledges and addresses the inherent complexities and subtleties in understanding how AI tools like ChatGPT influence user engagement across various knowledge-sharing platforms.

A key limitation of not pursuing a causal analysis is the presence of confounding variables that could influence the study's findings. These include changes in community guidelines, moderation policies, or external technological advancements concurrent with the introduction of ChatGPT, which could independently affect user engagement. Additionally, natural fluctuations in user behavior, such as seasonal activity changes or shifts in the demographic makeup of forum users, might also play a role in the observed trends. While these factors could impact the outcomes, our goal is not to isolate the effect of ChatGPT but rather to observe and document patterns of engagement that coincide with its introduction.

Despite these limitations, I focus on providing insights into the associations and patterns of user engagement with the introduction of ChatGPT. By documenting these trends, the research contributes to a broader understanding of how AI tools are integrated into and interact with digital knowledge ecosystems. This study serves as a foundational exploration, paving the way for future research that might employ more rigorous methods to ascertain causality and further delve into the dynamics of AI's influence on online communities.

6.2 Implications

The indirect influence of generative AI tools like ChatGPT on digital platforms such as Stack Exchange offers a unique perspective on their economic implications. While these AI tools are not directly integrated into these platforms, their availability and use by individuals can still profoundly impact the economic landscape of digital knowledge ecosystems.

The research results indicated significant variability in how different Stack Exchange

threads responded to the presence of ChatGPT. Threads like Stack Overflow and Code Review exhibited a substantial decrease in user interactions. This decline suggests that users may be bypassing traditional forum interactions in favor of quick answers from AI tools, potentially leading to reduced user-generated content. Economically, this could impact platforms' revenue models, which are often heavily dependent on active user engagement and sustained traffic for advertising revenue.

The positive coefficients observed in threads like Philosophy and Judaism, suggest that in some areas, while AI provides initial answers, the depth and complexity of certain topics drive users back to the forums for more thorough discussions. The variability in AI's impact across different domains, underscores the need for nuanced policy frameworks that can address potential issues such as data privacy, misinformation, and equitable access to AI tools. The concerns of copyrighted training data, and being substituted by generative AI has led platforms like Stack Overflow to create partnerships with generative AI companies [4]. Thus, continued partnerships and regulations are likely on way and needs to be adaptive and inclusive to help mitigate risks while fostering a competitive and innovative market environment.

6.3   Future Work

In terms of dataset expansion, an interesting area for future work would be analysis beyond Stack Exchange and looking at other open knowledge platforms like Reddit. Reddit has 70 million active users and 2.2 million subreddit threads[5], making it an ideal platform to investigate how generative AI might be shaping knowledge dissemination across different topics.

Further exploration in education research could offer more insight into how different changes in each thread affect users' ability to learn from these platforms. Finding the

---

[4]Stack Overflow, "OpenAI Partnership," accessed May 9, 2024, `https://stackoverflow.co/company/press/archive/openai-partnership`.

[5]Business of Apps, "Reddit Statistics," 2023, accessed April 30, 2024, `https://www.businessofapps.com/data/reddit-statistics/`.

right balance between utilizing open-source knowledge platforms like Stack Exchange and generative AI tools could enhance learning and deepen understanding across subjects (Chen et al., 2023). Future work should explore how to optimally integrate these technologies into different courses. The ideal approach for using generative AI in a history[6] or literature[7] class, for example, may differ from math or science. Investigating subject-specific strategies would allow students to engage with knowledge interactively while developing critical thinking skills.

Comprehensive training is also crucial to ensure that students and educators can harness these technologies effectively and responsibly. Such efforts can inform policies surrounding generative AI's use in education, given ongoing tensions[8]. By understanding how learners interact with diverse knowledge mediums like AI models, educators can develop methods that leverage their potential while addressing limitations.

7   Conclusion

In conclusion, this research provides valuable insights into how ChatGPT impacts user engagement across different knowledge domains within the Stack Exchange network. While the results did not reveal a straightforward pattern as anticipated, they highlight the complex ways in which generative AI interacts with human users across diverse forums.

Moreover, building upon the ideas by Marshall McLuhan in "The Medium is the Message," future research could delve deeper into the psychological and cognitive implications of how different mediums – be it open knowledge platforms, generative AI models, or a combination thereof – influence the way users engage with and assimilate knowledge. Conducting controlled experiments and user studies could provide valuable information on the nuanced

---

[6]"Simulating History with ChatGPT," Res Obscura, accessed April 30, 2024, `https://resobscura.substack.com/p/simulating-history-with-chatgpt`.

[7]Michael J. Shehane, "Because ChatGPT, I Can Now Enjoy Shakespeare," LinkedIn, 2023, `https://www.linkedin.com/pulse/because-chatgpt-i-can-now-enjoy-shakespeare-michael-j-shehane-diubc/`.

[8]M. Elsen-Rooney, "NYC education department blocks ChatGPT on school devices, networks," 2023, accessed January 24, 2023, `https://ny.chalkbeat.org/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence`.

interplay between the medium and the message, potentially forming the development of more effective and engaging knowledge sharing experiences.

Ultimately, as generative AI continues to evolve and permeate various aspects of society, it is crucial to maintain a holistic perspective on its impact. Although this study focused on quantifying the effects on user engagement within online knowledge ecosystems, the broader societal, ethical, and philosophical implications of these technologies warrant further investigation. Interdisciplinary collaborations between researchers, policymakers, and diverse stakeholders could help ensure that the potential of generative AI is harnessed responsibly and equitably, fostering a future where human knowledge and understanding are enriched, rather than diminished, by technological advancements.

8 References

Austin, J., A. Odena, M. I. Nye, M. Bosma, H. Michalewski, D. Dohan, E. Jiang, C. J. Cai, M. Terry, Q. V. Le, and C. Sutton (2021). Program synthesis with large language models. *CoRR abs/2108.07732*.

Brand, J., A. Israeli, and D. Ngwe (2023, March). Using gpt for market research. (23-062). Available at SSRN: https://ssrn.com/abstract=4395751 or http://dx.doi.org/10.2139/ssrn.4395751.

Burtch, G., D. Lee, and Z. Chen (2023). The consequences of generative ai for ugc and online community engagement. *SSRN Electronic Journal*.

Chen, B., X. Zhu, and F. Díaz del Castillo H. (2023). Integrating generative ai in knowledge building. *Computers and Education: Artificial Intelligence*.

Chen, M., J. Tworek, H. Jun, et al. (2021). Evaluating large language models trained on code. *CoRR abs/2107.03374*.

del Rio-Chanona, M., N. Laurentsyeva, and J. Wachs (2023). Are large language models a threat to digital public goods? evidence from activity on stack overflow. *arXiv preprint arXiv:2307.07367*.

Fabriz, S., J. Mendzheritskaya, and S. Stehle (2021). Impact of synchronous and asynchronous settings of online teaching and learning in higher education on students' learning experience during covid-19. *Frontiers in Psychology 12*.

Felten, E., M. Raj, and R. Seamans (2021, December). Occupational, industry, and geographic exposure to artificial intelligence: A novel dataset and its potential uses. *Strategic Management Journal 42*(12), 2195–2217.

Gans, J. S. (2024, April). Copyright policy options for generative artificial intelligence. *SSRN Electronic Journal*.

McLuhan, M. (1994). *The Medium Is the Message.* MIT Press.

Nonaka, I. and H. Takeuchi (1995). *The Knowledge-Creating Company: How Japanese Companies Create the Dynamics of Innovation.* Oxford University Press.

Polanyi, M. (1966). *The Tacit Dimension.* University of Chicago Press.

Wang, Z., Y. Pang, and Y. Lin (2023). Large language models are zero-shot text classifiers. *arXiv preprint arXiv:2312.01044*.

White, J., Q. Fu, S. Hays, M. Sandborn, C. Olea, H. Gilbert, A. Elnashar, J. Spencer-Smith, and D. C. Schmidt (2023). A prompt pattern catalog to enhance prompt engineering with chatgpt. *arXiv preprint arXiv:2302.11382*.

Zhou, E. and D. Lee (2024, March). Generative ai, human creativity, and art. *PNAS Nexus 3*, pgae052. Available at SSRN: https://ssrn.com/abstract=4594824 or http://dx.doi.org/10.2139/ssrn.4594824.

9 Appendix

Table 5: Summary Statistics Question Count per Week by Thread (Part 1)

Thread	Mean	Std Dev	Min	Max
Academia	58.70	13.30	29	97
Android	43.50	10.80	22	86
Apple	118.67	25.14	62	187
AskUbuntu	380.13	80.77	197	617
Blender	246.54	39.89	179	371
Chemistry	50.48	19.09	24	123
Christianity	24.39	7.47	11	62
CodeGolf	12.89	6.32	3	39
CodeReview	56.10	19.58	20	108
Computer-Science	61.48	21.31	26	144
DBA	107.11	21.86	50	173
DIY	155.90	33.52	102	283
Drupal	30.50	10.59	10	64
Electronics	306.46	47.70	191	473
English	79.15	26.71	36	170
English Language Learners	152.34	52.58	65	321
Gaming	46.27	20.04	17	107
Hermeneutics	35.12	15.31	9	79
Judaism	45.43	16.55	17	100

Table 6: Summary Statistics Question Count per Week by Thread (Part 2)

Thread	Mean	Std Dev	Min	Max
Magento	71.99	25.89	24	167
Meta StackExchange	36.61	12.30	14	107
Meta StackOverflow	35.38	11.65	13	71
Money	43.14	20.38	16	166
Philosophy	38.54	12.83	16	85
Physics	365.91	66.30	250	558
Politics	31.08	12.83	11	93
Salesforce	160.42	34.21	73	243
SciFi	77.90	13.24	50	113
ServerFault	191.72	36.55	107	310
SharePoint	47.13	16.54	18	97
Software Engineering	32.29	11.80	12	75
StackOverflow	24223.01	5551.60	9903	34030
Stats	296.45	45.87	128	429
TeX	302.96	56.32	152	452
User Experience	17.96	4.98	7	33
WordPress	71.25	21.97	33	151
Worldbuilding	54.20	21.71	17	113