PPL Institute
Psychology, Politics, Liberty

Who Are You Talking To? The Discernment of AI and Human Content on Social Media

Ryan R. Rajpal
Director

Alexis R. Grayon
Nonresident Fellow

Paige Iino
Research Assistant

The digital space and social media have become deeply integrated with society and the day-to-day operations of billions of individuals around the globe. The internet is commonly viewed as a way to connect people from across the world easily. The introduction of social bots and artificial intelligence (AI) to the social media landscape contributes to what can be called the "dead internet theory," which speculates that the internet is dominated primarily by bot accounts, not people. This theory indicates a paradigm shift in how social media and its influence are studied. While social media has been found to influence public opinion and attitudes, it is not entirely clear whether social media users can effectively differentiate between AI-generated and human-made content on social media. As discernment between AI and human content may affect the extent to which social media can influence opinions and attitudes, this study deploys an original survey to measure participants' ability to differentiate between AI and human posts accurately. Participants were more likely to identify social media posts as AI-generated, but were largely able to differentiate between AI and human content successfully. ChatGPT was able to successfully identify the creator of each post, demonstrating the potential of AI to moderate content on social media successfully.

Keywords: Artificial intelligence, Social media, Public opinion, Politics, Communication

Published: 11 July 2025

Introduction

Since its conception in the 1990s, social media has been viewed as a way for people to connect and communicate with others around the world (Dorsey et al., 2020). In this way, the internet has created a digital reality parallel to our physical reality, including spaces for entertainment, information gathering, socialization, and commerce. In its early form, the internet was untouched by corporations and governmental organizations, existing solely for and from the people (Gerbaudo, 2017). However, the internet has since undergone significant commodification and now hosts political actors and corporations that create content targeting users for specific purposes, like mass influence over opinion and behavior. In 2008, Barack Obama notably used social media in his campaign to mobilize voters and was a pioneer of this campaign method (Aaker & Chang, 2009). Since then, social media campaigns have become the norm in political campaigning. Similarly, corporations utilize social media not only to advertise products and services but also to conduct market research (Patino et al., 2012). Thus, although social media is still for the people, it is no longer solely for the people.

While the internet underwent commodification during the 2010s, the dead internet theory identifies a new era for social media, one characterized by bot and artificial intelligence (AI) accounts (Ghantous, 2025; Walter, 2024). The theory holds that the internet consists largely of bot and AI accounts that create posts intended to sway public opinion on a variety of subjects. Previously understood as a conspiracy theory, there is growing evidence that transformations aligned with the dead internet theory are occurring, especially as AI continues to be refined and increasingly mimics human activity (Fetzer, 2024; Pocol, 2025). In sum, these accounts and their content are becoming harder to detect. As such, the authenticity of perceived human-to-human interactions becomes questionable in the context of posts and comments on various platforms

that are largely text-based, such as Reddit, X (formerly known as Twitter), and even Instagram. While most social media platforms have countermeasures against bot and AI accounts, these accounts persist and post content regularly (Orabi et al., 2020; Hajli et al., 2022). Thus, we investigate the degree to which people are able to differentiate between three AI-generated and three human-made social media posts, or a total of six conditions, fashioned after Reddit posts. ChatGPT was also tasked with identifying the creator of the six conditions. We find a discrepancy between the ability of humans and AI to identify the creator of each post. Before concluding, we discuss possible benefits and harms associated with the refinement of AI in the context of social media.

People and Social Media

Social media platforms are used largely by younger generations, such as Generation Z (Gen-Z), born between the years 1997-2012, and millennials, born between 1981-1996, and have been attributed to easy access to information, entertainment, and socialization (Dimock, 2019). Among these platforms are influencers, who are self-motivated and use social media platforms as a way to connect with others. They play a prominent role in shaping users' opinions, behavior, and consumer habits, impacting the consumer market regarding clothes, shoes, cosmetics, and services. However, they also have an impact on other topics such as public health and political opinion (Hogsnes et al., 2024; Stocking et al., 2024). Gambo and Özad (2020) find that media habits differ between men and women, such that women pursue specific content areas while men browse content more generally. As a result, women may seek out influencers who touch on certain topics more often than men, who, at large, are passively influenced by social media algorithms.

Related, political influencers often provide easy synthesis of complex information for the public, making information easier to access (Goodwin et al., 2023). Although many are self-motivated, some influencers are contacted by political organizers. Regardless of self-motivation or political coordination, influencers are more inclined towards cause-based campaigns. Political organizations that enlist the help of influencers to spread a message actively partake in astroturfing, or the orchestration of spreading a message to influence public opinion (2023). Thus, while social media is used as a way for individuals to access information and communicate, it is also a space where organizations actively try to sway opinion en masse.

Social media users also use these platforms for news consumption; however, Shu et al. (2018) find that the quality of the information is lower than that of traditional news outlets, especially given the prevalence of fake news. Fake news is shared by a few users on social media platforms, but real news is shared by even fewer users. By constructing two datasets with news content and social context information, Shu et al. (2018) find that verified users are more likely to trust real news than unverified users. Both bot and AI accounts are largely unverified on social media, as the verification process on various platforms often includes checks that the account is owned by real people or organizations. However, on some social media such as X, verification can be bought, which introduces ambiguity regarding the nature of the account (human, bot, or AI). The authors also find that those who publish fewer posts are more likely to trust real news than those who publish more posts. It is worth noting that bot and AI accounts post at much higher frequencies than humans, contributing to the fake news in circulation on social media (Javed et al., 2025).

There are countermeasures in place to inform users about fake news on social media. Social media like X and Reddit have community notes or moderators, often informed on a

subject, to provide a disclaimer or flag any factually incorrect information. In an electroencephalogram (EEG) experiment, Moravec et al. (2019) measured brain activity while participants were tested on detecting fake news on social media of 50 fact-based headlines. Flags for fake news triggered greater cognitive activity but ultimately did not influence the participants' beliefs. Instead, confirmation bias persisted, demonstrating that users were more likely to believe headlines aligned with their political beliefs than those that challenged their opinions. This signals poorly for the incorporation of AI on social media, which can be used to purposefully spread convincing fake news, and leads to the question fundamental to this research:

RQ: Can people differentiate between AI-generated and human-made social media posts?

Artificial Intelligence and Social Media

Bots and AI accounts publish and repost social media content at greater frequencies than human accounts, which can contribute to the disproportionality between fake and real news on social media. Increasing accessibility and higher quality of content generated by AI have made it more difficult to distinguish AI-generated content from human-made content on social media platforms like Instagram, X, and Reddit. Yang et al. (2024) use a dataset of 1,420 Twitter accounts and find that 0.021% - 0.044% of accounts use profile pictures of AI-generated faces (8k-17k daily active accounts of 40 million users) and are responsible for 0.022% - 0.053% of tweets. These AI accounts can spread misinformation, but beyond that, their being on social media can lead to inauthenticity online, degrading the integrity of the platforms as a way for people to connect. These accounts use ChatGPT to create human-like content, often with the purpose of impersonation, scamming, spamming, and astroturfing.

Furthermore, most AI accounts on Twitter were not verified. These accounts also may have AI-generated profile pictures, which can be determined by defects such as clothing imperfections in collars and hats. Many AI companies like OpenAI and Midjourney do not use Generative Adversarial Networks (GANs), which generate images based on previous data, for image generation; they use more complex models that produce better images (2024). As such, intervention strategies are increasingly necessary to regulate generative AI models and inauthentic accounts.

Park et al. (2024) use Instagram to conduct a user study on how people perceive and evaluate AI-generated art content and accounts compared to those created by humans. Previous work has already indicated that people find it challenging to distinguish between AI and human content (Fetzer, 2024; Pocol, 2025). Building on this, the authors found that user perceptions and evaluations significantly differed across AI, influencer, and public accounts. Participants tended to perceive both AI and influencer accounts as more attractive and credible than public accounts, making them more likely to follow or recommend them. Post-study interviews revealed that many participants believed the AI and influencer accounts belonged to experts with a keen sense of art, whereas the public accounts were perceived as belonging to art enthusiasts or amateur artists. As such, we believe that:

H1a: Participants will correctly identify the source of a majority of individual posts.
H1b: Few participants will correctly classify all six posts.

Nonetheless, participants reported using several mechanisms to identify AI-generated content. Common indicators included inconsistent or skewed imagery, excessive uniqueness, oddly consistent video speed, and inconsistent or illegible text. Other tell-tale signs were a lack

of detail, zoom-in and fade-out effects, and a distorted or unrealistic representation of subjects. In textual descriptions, participants noted that perfect grammar, number of hashtags, and basic emoji usage were patterns that pointed to AI-generated content. The study's findings confirm the growing difficulty in distinguishing AI-generated content from human-made content.

However, AI and machine learning can fulfill the role of community contributors or moderators, too, by detecting and removing fake news (Aïmeur et al., 2023). Hajli et al. (2022) find that Actor Network Theory (ANT) provides a method for understanding human and non-human relations on social media, specifically on Twitter. With applications to person-to-person, person-to-object, and object-to-object relations, researchers use ANT to create an algorithm that sorts through the content and identifies bot-originated content. Using a dataset of 30,000 tweets, they use the algorithm to decipher which items are of bot origin with around a 77% accuracy rate. Deep learning for AI includes recurrent neural networks that provide the AI with "memory" and the ability to classify sound and signals— this is their proposed method for filtration on social media. Similarly, Alnagi et al. (2024) use the SHapley Additive exPlanations (SHAP) technique, an explanatory algorithm, for AI detection on social media and report around a 90% accuracy in their filtration. Given the increasing sophistication of AI in generating convincing human-like content, we also test ChatGPT's ability to accurately identify the creator of each post.

H2: ChatGPT will accurately identify the source (AI vs. human) of all six social media posts.

Methods

To address the research question, a survey was developed to measure participants' ability to differentiate between AI-generated and human-made social media posts. The survey was

designed using Google Forms and made publicly available on the PPL Institute website under the "Research" page. Participants were recruited through the researchers' networks on Instagram and LinkedIn, resulting in a sample of 34 participants. All participants were members of Gen-Z.

To create the experimental conditions, the researchers used ChatGPT (GPT-4-Turbo) to generate three text-based social media posts. The prompt provided to ChatGPT was: "Provide me three human-like social media posts for a text-based social media platform like Reddit and Twitter. Of the three posts, one should be about politics, one should be about pop culture, and one should be a meme." For the memetic post, ChatGPT provided a description of an image. Researchers then prompted ChatGPT again to generate an image based on its description. Human-made posts were written by the researchers on the same topics as the AI-generated content. Additionally, researchers created a meme using a popular meme format referencing the 2024 Olympics shooting event. Thus, there were six conditions: r/politicaldiscussion (AI and Human), r/television (AI and Human), and r/memes (AI and Human).

Figure 1
AI Post: r/politicaldiscussion

r/politicaldiscussion

Title: Why does it feel like we're stuck in a political loop every 4 years?

Text: I've been thinking lately about how each election cycle feels like a rerun of the last — the same debates, the same wedge issues, the same outrage. Doesn't matter which party is in power, real progress feels glacial. Is this just the nature of a two-party system, or are we genuinely stuck in a kind of performative democracy? Would love to hear perspectives from people outside the U.S. too — is this just an American thing or more universal?

(OpenAI, 2025)

Figure 2
Human Post: r/politicaldiscussion

Title: Each election is the same…
Text: It just seems like each election leaves Americans disappointed in either the president elect or the people that elected them, regardless of party. No one is ever satisfied with the policies and I'm beginning to wonder if that's just how it is in a two party system. I don't know, is that why Europeans are happier?

Figure 3
AI Post: r/television

Title: Is anyone else lowkey burned out by all the reboots?
Text: Not trying to be a hater, but between the Office reboot, yet another Harry Potter series, and every '90s movie getting a "legacy sequel," I'm feeling more exhausted than excited. I get the nostalgia factor, but does it feel like creativity in mainstream media is running on empty? Or are we just in a weird phase before something fresh breaks through again?

(OpenAI, 2025)

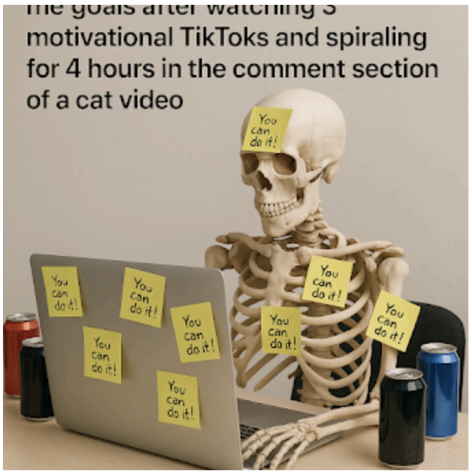
Figure 4
Human Post: r/television

Title: What TV show could have done without a reboot?
Text: I'm talking about stuff that is objectively good as it stands. The Harry Potter tv show seems unnecessary as the movies are highly acclaimed. It just feels like they're chasing the same success. Same thing with the Office spinoff The Paper, I guess you could argue a spinoff is a fresh take but that seems like a stretch to me.

Figure 5
AI Post: r/memes

r/memes

Me trying to be productive for once...



(OpenAI, 2025)

Figure 6
Human Post: r/memes

r/memes

Fake compared to real productivity

The presentation of posts within the survey was randomized, and all participants were shown the same set of posts in randomized order, including the conditions with images. The full survey is presented in the appendix. Each post was presented on its own page, on which participants were asked to determine whether the post was generated by AI or created by a human. To measure AI's ability to differentiate between human-made and AI-generated posts, each post was uploaded to ChatGPT (GPT-4-turbo), including the images used in the two r/memes posts. Participants were asked to classify each condition one at a time as either AI- or human-generated. The dataset was cleaned and analyzed using R version 4.4.2 (R Core Team, 2024). Paired t-tests were used to test for statistically significant differences between AI and human conditions. All tests were two-tailed with a significance threshold set at $p < 0.05$. Cohen's d for paired samples was calculated to assess the effect size of the difference in participants' ability to classify AI-generated versus human-made posts.

Results

This study consists of 34 participants situated firmly in Gen-Z, aged 18-27 years old. Of the 34 participants, 16 were male, 16 were female, and two identified as other. Social media usage was generally high among the sample. When coded as values from 1 to 7 based on a 7-point Likert scale, the average social media usage score was 3.94, between "Somewhat infrequently" and "Moderate frequency" (Table 1). About 67% of participants reported seeing AI content often online, and the average participant believed that 30% of accounts online were AI, as indicated by the average belief score in Table 1. With regard to AI usage, 85% of participants reported generative use of AI for text, while only 35% of participants used AI to create images. Notably, 15% of participants had never used generative AI at all. Attitudes concerning AI

accounts posting online were largely negative, as 85% of participants believed it was bad for such accounts to post on social media. Similarly, 85% of participants believed that the government should take a stricter approach to AI regulation. The descriptive statistics for exploratory opinion variables are summarized in Table 1.

Table 1
Select Variable Descriptions

Variable	Response	Percent of Responses
Social Media Usage	Very infrequently	5.9%
	Infrequently	5.9%
	Somewhat infrequently	5.9%
	Moderate frequency	2.9%
	Somewhat frequently	20.6%
	Frequently	26.5%
	Very frequently	32.4%
Generative AI Usage	Never	20.6%
	Rarely	26.5%
	Sometimes	26.5%
	Often	26.5%
	Always	0.0%
Attitude Towards AI on Social Media	Favorable	85.3%
	Unfavorable	5.9%
	Unsure	8.8%
Attitude Towards Stricter Government Regulation	Favorable	85.9%
	Unfavorable	2.9%
	Unsure	11.8%
Believed Percentage of Online AI Accounts (Median, IQR)		30% (20-30%)

The analysis is based on two paired t-tests. The first compares participants' ability to differentiate between AI-generated posts and human-made posts. The results from this first test show that participants were able to distinguish between the two: AI-generated posts were distinguished as made by AI while human posts were seen as human-made $t_{33} = -4.37$ ($p < 0.001$) with a mean difference of -0.304 and a 95% confidence interval of (-0.445, -0.163). There was a large effect, with participants substantially more accurate at identifying AI-generated content compared to human-made content, d = -1.01 (95% CI: -1.74, -0.46). This supports H1a; thus, we reject the null hypothesis that people would be unable to identify the origin of most posts.

The second test focused specifically on text-only posts. Participants were slightly more accurate at identifying AI-generated content compared to human-made content, with a mean difference of -0.103 (95% CI: -0.308, 0.102) and t-value of -1.02 (p = 0.31). However, this difference was not statistically significant. The effect size was small, d = -0.28 (95% CI: -0.83, 0.28), suggesting a modest difference in participants' ability to differentiate between the text-based content. These results indicate that participants were less successful in distinguishing between AI and human posts when evaluating text-only content, as illustrated by Figure 7. Figure 8 demonstrates that most participants (n = 22) were able to successfully identify at least four of the six posts correctly, while only five participants successfully identified each post, supporting H1b. Thus, we reject the null hypothesis that most or none of the participants would be able to successfully identify each post.

Figure 7
Differences in Differentiation Across Conditions

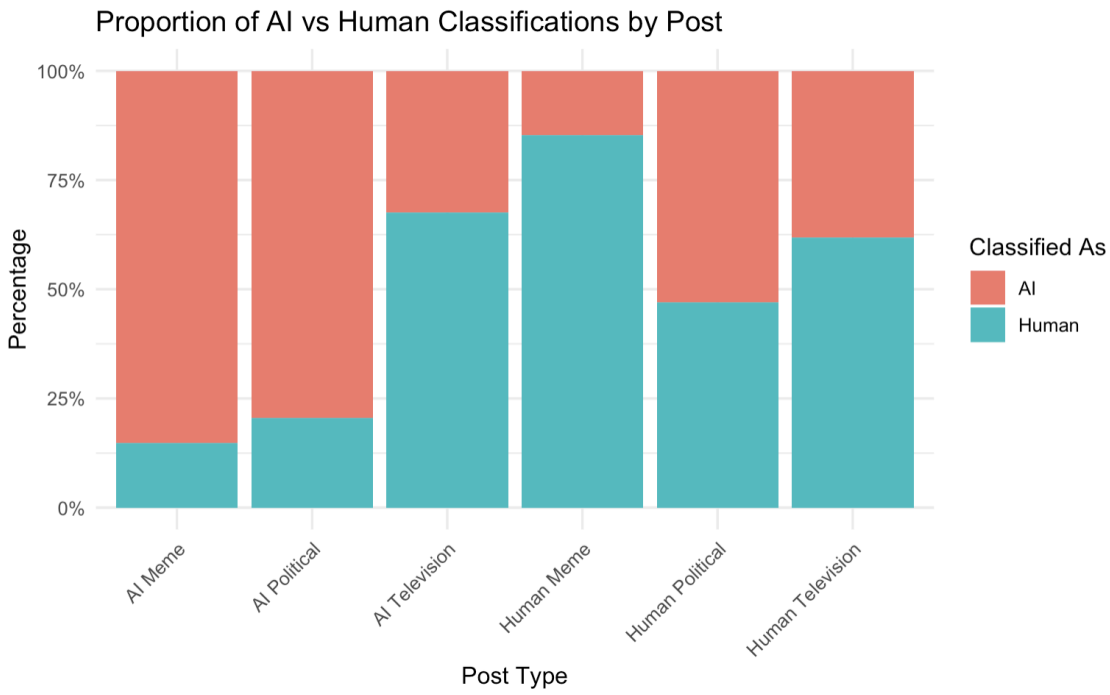
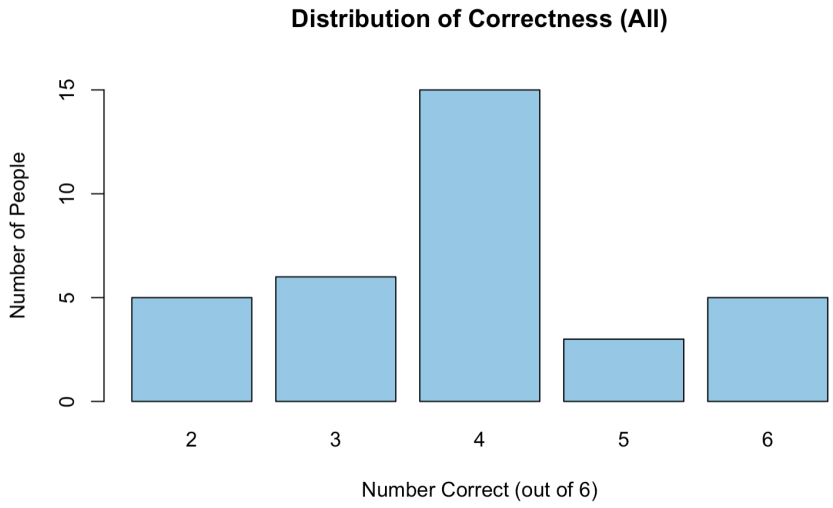


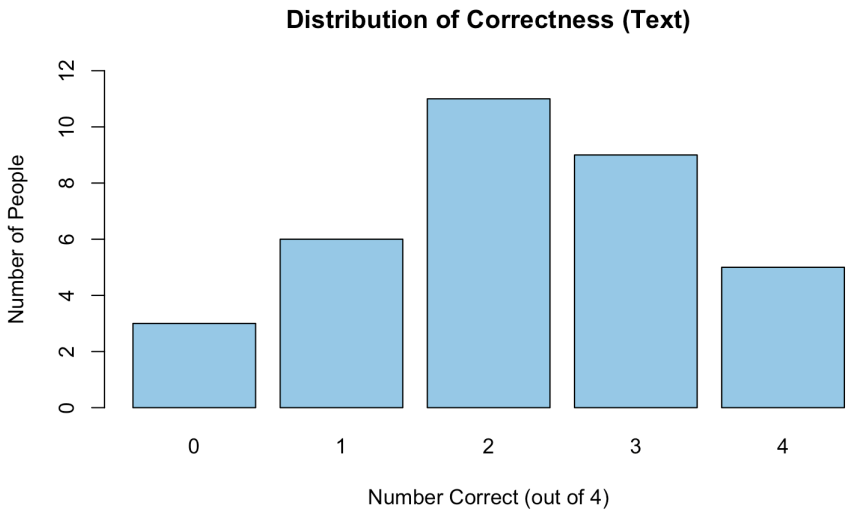
Figure 8
Correct Responses Across All Social Media Posts



To understand participants' ability to differentiate between solely text-based AI and human content, we removed the posts with images from the bar plot in Figure 9. Of the 34 participants, 59% (n = 20) were unsuccessful at differentiating between more than half of the

text-based content, and only 15% (n = 5) were able to discern the creator of each post successfully.

Figure 9
Correct Responses Across Text-Based Social Media Posts



ChatGPT was prompted to identify the six social media posts as either AI-generated or human-made and used parameters established in tone and text style. Through these parameters, ChatGPT correctly identified the creator of all six of the social media posts as seen in Figure 10, supporting H2. Thus, we can reject the null hypothesis that ChatGPT would not be able to accurately identify the creator of each post.

Figure 10
ChatGPT's Ability to Differentiate

Subreddit	Title (Shortened)	Verdict
r/memes	"me trying to be productive for once..."	**Human-made**
r/memes	"fake vs real productivity"	**AI-made**
r/politicaldiscussion	"Why does it feel like we're stuck..."	**AI-made**
r/politicaldiscussion	"Each election is the same..."	**Human-made**
r/television	"Is anyone else lowkey burned out..."	**AI-made**
r/television	"What TV show could have done without a reboot?"	**Human-made**

(OpenAI, 2025)

Discussion

This study sought to understand how well people are able to differentiate between AI and human content on text-based social media platforms such as Reddit. The findings show that the Gen-Z participants use social media to a large extent and that 85% of participants use AI like ChatGPT to generate text. However, fewer individuals use AI beyond text generation for purposes like image, video, or audio generation. Given this text-based use of AI, it was expected that participants would be able to accurately differentiate between AI and human content with relative ease. The findings partially support this theory, as 65% of participants were able to accurately parse between four or more of the six social media posts. However, 59% of the sample failed to correctly identify more than half of the text-based posts, raising questions about whether their success may have been due to chance. This discrepancy may be attributed to the ease with which AI can mimic human text or the inattentiveness of participants to nuanced differences between AI and human text.

Participants' relative ease in discerning the creator of each post for the r/memes condition may also be attributed to clear differences in the images that AI can generate compared to those of humans. That is, there are overly smooth qualities or shadowing inconsistencies in AI-generated images that make the content easier to discern. Thus, when the image-based conditions were removed from the analysis, such that only data regarding text-based posts were present, participants had less success differentiating between the AI and human content. On the other hand, ChatGPT successfully identified the creator of each condition based on differences in the language— AI-generated text having extreme levels of conciseness compared to the more unpolished nature of human-written text.

While it appears that people are able to largely differentiate between the content of AI and humans, the data spotlights a subgroup of individuals who have difficulty with higher-order discernment. Thus, the effectiveness of ChatGPT in differentiating between AI and human content indicates that AI can be implemented to flag and remove non-human content on social media platforms that have rules against AI accounts. However, it is unclear what future directions platforms may take regarding the regulation of human accounts that repost AI content, such as AI art, if any.

This research is limited in that there is a small sample size of only 34 participants. For greater representativeness of social media users, future studies based on similar research questions should obtain a larger number of participants. The small sample size also inhibits the ability to meaningfully run more complex analyses, such as regressions with a number of exploratory variables that could provide a more nuanced understanding of the dynamics that influence a person's ability to differentiate between AI and human content, such as social media usage and AI usage. Further, the sample is unrepresentative of the population in that participants

opted into the study and were recruited from the researcher's personal networks (Instagram story, LinkedIn post, in-person recruitment). Thus, a formal survey distribution platform should be used in future research for a more representative and generalizable sample.

Still, this study demonstrates that as AI is refined and able to generate more human-like content, the ability to differentiate between AI and human content may further dissipate. Conversing with other people is a way for someone to better understand their own opinions or change them. People who believe they are interacting with a person may experience changes in opinion or behavior as a result of the interaction. Interacting with AI may lead to similar changes, but it would be deceitful and unrepresentative of a community of people. As such, social media platforms may be keen to use AI to detect AI content effectively and remove it, thus stalling or even reversing the dead internet. Social media platforms may feel inclined to implement such measures themselves; however, the majority of participants indicated that they believed the government should intervene to help regulate AI.

Conclusion

In conclusion, this research sought to understand and measure the human ability to differentiate between AI and human content on social media, backdropped by the rise of AI accounts on these platforms and the further training of AI on human content. In particular, this study uses content fashioned after Reddit posts and found that in posts with images, people were largely able to tell AI content apart from human content. Although participants reported frequent social media usage and seeing AI content online often, participants were not able to easily distinguish between AI and human text-based content.

As AI continues to be developed and implemented in spaces like social media, it is pertinent that researchers continue to investigate how AI relates to and interacts with people. While people can differentiate between AI and human content more often than not, platforms may benefit from implementing measures to reduce the amount of AI content shared, as it would provide an authentic space for individuals to engage with others. Government intervention regarding AI may also be viewed favorably by social media users. This research serves as an early contribution to the evolving conversation about AI's role in shaping online communication and signals a need for precaution when navigating the digital landscape.

References

Aïmeur, E., Amri, S. & Brassard, G. Fake news, disinformation and misinformation in social media: a review. *Soc. Netw. Anal. Min*. 13, 30 (2023). https://doi.org/10.1007/s13278-023-01028-5

Alnagi, E., Ahmad, A., Al-Haija, Q. A., & Aref, A. (2024). Unmasking fake social network accounts with explainable intelligence. *International Journal of Advanced Computer Science and Applications*, *15*(3). https://doi.org/10.14569/ijacsa.2024.01503125

Dimock, M. (2019). Defining generations: Where Millennials end and Generation Z begins. *Pew Research Center*. Retrieved June 23, 2025, from https://www.pewresearch.org/short-reads/2019/01/17/where-millennials-end-and-generation-z-begins/

Dorsey, J., Williams, E., & Stone, B. (2020, May 20). *The Evolution of Social Media: How Did It Begin, and Where Could It Go Next?*. Maryville University Online. Retrieved July 2, 2025, from https://online.maryville.edu/blog/evolution-social-media/

Fetzer, M. (2024, May 14). *Q&A: The increasing difficulty of detecting AI-versus human-generated text*. Information Sciences and Technology. https://www.psu.edu/news/information-sciences-and-technology/story/qa-increasing-difficulty-detecting-ai-versus-human

Gambo, S., & Özad, B. O. (2020). The demographics of computer-mediated communication: A review of social media demographic trends among social networking site giants. *Computers in Human Behavior Reports*, 2, 100016. https://doi.org/10.1016/j.chbr.2020.100016

Gerbaudo, P. (2017). From cyber-autonomism to cyber-populism: An ideological analysis of the evolution of Digital Activism. *TripleC*, *15*(2), 477–489. https://doi.org/10.31269/triplec.v15i2.773

Ghantous, Danny (2025) *Between the Self and Signal: The Dead Internet & a Crisis of Perception.* [MRP]. https://openresearch.ocadu.ca/id/eprint/4676

Goodwin, A., Joseff, K., Riedl, M., Lukito, J., & Woolley, S. (2023). Political Relational Influencers: The Mobilization of Social Media Influencers in the Political Arena. *International Journal Of Communication*, 17, 21. https://ijoc.org/index.php/ijoc/article/view/18987

Hajli, N., Saeed, U., Tajvidi, M., & Shirazi, F. (2022). Social Bots and the Spread of Disinformation in Social Media: The Challenges of Artificial Intelligence. *Brit J Manage*, 33: 1238–1253. https://doi.org/10.1111/1467-8551.12554

Hogsnes, M., Grønli, T. M., & Hansen, K. (2024). Exploring Influencers' Commercial Content on Instagram. *Journal of Interactive Advertising*, 24(2), 156–168. https://doi.org/10.1080/15252019.2024.2316114

Javed, D., Jhanjhi, N. Z., Khan, N. A., Ray, S. K., Al-Dhaqm, A., & Kebande, V. R. (2025). Identification of Spambots and Fake Followers on Social Network via Interpretable AI-Based Machine Learning. *IEEE Access*, 13. https://doi.org/10.1109/ACCESS.2025.3551993

Jennings, F. J., Suzuki, V. P., & Hubbard, A. (2020). Social Media and Democracy: Fostering Political Deliberation and Participation. *Western Journal of Communication*, *85*(2), 147–167. https://doi.org/10.1080/10570314.2020.1728369

Moravec, P. L., Minas, R. K., & Dennis, A. R. (2019). Fake News on Social Media: People Believe What They Want to Believe When it Makes No Sense At All. *MIS Quarterly*, 43(4), 1343–1360, A1–A13. https://www.jstor.org/stable/26848107

OpenAI. (2025). *ChatGPT (GPT-4 Turbo)* [Large language model]. https://chat.openai.com/

Orabi, M., Mouheb, D., Al Aghbari, Z., & Kamel, I. (2020). Detection of bots in social media: a systematic review. *Information Processing & Management*, 57(4), 102250. https://doi.org/10.1016/j.ipm.2020.102250

Park, J., Oh, C., & Kim, H. Y. (2024). AI vs. Human-generated content and accounts on Instagram: User preferences, evaluations, and ethical considerations. *Technology in Society*, 79, 102705. https://doi.org/10.1016/j.techsoc.2024.102705

Patino, A., Pitta, D. A., & Quinones, R. (2012). Social media's emerging importance in market research. *Journal of consumer marketing*, 29(3), 233-237. https://doi.org/10.1108/07363761211221800

Pocol, A. (2025, March 25). *As generative AI becomes more sophisticated, it's harder to distinguish the real from the deepfake*. The Conversation. https://theconversation.com/as-generative-ai-becomes-more-sophisticated-its-harder-to-distinguish-the-real-from-the-deepfake-225768

R Core Team (2024). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Stocking, G., Wang, L., Lipka, M., Matsa, K. E., Widjaya, R., Tomasik, E., & Liedke, J. (2024, November 18). *Influencers Who Often Post About News, and Who Sees Their Content.* Pew Research Center. Retrieved July 7, 2025, from https://www.pewresearch.org/journalism/2024/11/18/americas-news-influencers/

Shu, K., Wang, S., & Liu, H. (2018). Understanding user profiles on social media for fake news detection. *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)* (pp. 30–435). https://doi.org/10.1109/mipr.2018.00092

Walter, Y. (2024). Artificial Influencers and the dead internet theory. *AI & SOCIETY*, *40*(1), 239–240. https://doi.org/10.1007/s00146-023-01857-0

Yang, K. C., Singh, D., & Menczer, F. (2024). Characteristics and prevalence of fake social media profiles with AI-generated faces. *arXiv preprint arXiv:2401.02627*. https://arxiv.org/abs/2401.02627