

# Generative AI as a Dangerous New Form of Media

Louis ROSENBERG

Unanimous AI

Pismo Beach, CA 93448, USA

## ABSTRACT

When policymakers and regulators assess the near-term risks of generative AI, they often focus on the dangerous ability to create and disseminate traditional misinformation at scale. This paper argues that we must expand this view and consider generative AI not merely as a vehicle for producing traditional content at scale, but as enabling an entirely new form of media that is adaptive and interactive in real-time and can be personalized for individual users. Referred to herein as Interactive Generative Media (IGM), this new form of targeted influence could be used to manipulate individual users at scale and compromise human agency. Two specific tactics are presented: Targeted Generative Advertising and Targeted Conversational Influence. The risks of each are described along with recommended policy protections.

**Keywords:** Generative AI, Large Language Models, Chatbots, Interactive Generative Media, Conversational Advertising, HCI.

## 1. INTRODUCTION

As policymakers confront the dangers of generative AI, they often focus on its ability to create and disseminate traditional misinformation at scale. This is because generative systems can now fabricate everything from images and videos, to articles, legal briefs, and scientific papers with convincing results. For example, recent research found that users cannot distinguish between tweets generated by GPT-3 and those created by real human users [1]. The risk, therefore, is that bad actors will flood the digital world with untruths, overwhelming our ability to distinguish the factual from the fabricated.

On the other hand, the dissemination of false and misleading content is already a significant problem online whether generated by humans or not. From this perspective, current methods for combatting traditional misinformation may be effective against this type of generative content. Regulators, therefore, must also focus on how generative AI could create entirely new threats that are not addressed by current protections. In particular, we must consider how generative AI technology can enable new forms of targeted content that are personalized, interactive and adaptive in real-time, and can be deployed at scale to target individual users on a person-by-person basis [2,3].

Referred to herein as *Interactive Generative Media*, this new form of personalized content has the potential to not only spread misinformation, but it could also threaten human agency through real-time deception, coercion, or manipulation. The phrase “epistemic agency” refers to an individual’s control over his or her own personal beliefs [4]. When epistemic agency is compromised by new forms of media, the political establishment can undermine democratic institutions by deploying propaganda, disinformation, and misinformation that supports authoritarian objectives, interests, or policies [5-8].

Mass media techniques have been used for generations to weaken agency in populations, but over the last decade this problem has been amplified by social media and its ability to

spread targeted influence with network effects [9]. This has been shown to drive polarization, promote radicalization, and foster discontent [10,11]. In this context, all new forms of media should be evaluated in their capacity for abuse and misuse, especially when it has the potential to threaten human agency [12].

## 2. INTERACTIVE GENERATIVE MEDIA

**Interactive Generative Media** (IGM) is informational content that is custom crafted by generative AI systems to maximize impact on individually targeted users and can adapt its persuasive tactics in response to real-time user reactions. Unlike traditional forms of targeted media (e.g., essays, articles, videos, memes, and posts) that are aimed at specific demographic groups, like buckshot fired at flocks of birds, generative AI will enable targeted content to be deployed more like heat-seeking missiles, zeroing-in on individual users for optimal effect.

In this way, Interactive Generative Media can be defined as personalized, adaptive, and interactive content that is deployed in real-time to maximize engagement and optimize influence on individually targeted users. It is no coincidence that driving user engagement and selling user influence are core business goals of most ad-based social platforms. It is therefore likely that IGM will be used aggressively by major corporations unless regulators put guardrails in place. In fact, Google, Meta, and Microsoft each announced plans in early 2023 to use generative AI to automate the creation of targeted ads [13, 14, 15]. If meaningful regulatory protections are not put in place, it is likely that two dangerous forms of interactive media will become standard practice among large technology platforms: “*Targeted Generative Advertising*” and “*Targeted Conversational Influence*.”

**Targeted Generative Advertising** can be defined as text-based, image-based, and video-based forms of informational content that look and feel like traditional advertising but is personalized by generative AI for maximum impact on individually targeted users. These ads will be created in real-time based on influence objectives provided by third-party sponsors that describe the advertising concept and messaging goals. This will be used in combination with personal background data accessed for the specific user for whom the custom content is being created. The personal data may include the user’s age, gender, interests, values, education-level, aesthetic and intellectual sensibilities, purchasing tendencies, and cultural biases.

In response to the influence objectives and personal data, the generative AI system will create the ad, customizing the layout, feature images, and language used to maximize effectiveness on that user. Everything down to the colors, fonts, and punctuation could be personalized along with age, race, and clothing styles of any persons shown in the imagery. And because tech platforms can track user engagement, the system will learn which promotional features work best on individual consumers, discovering for example, whether you respond better to redheads or brunettes – whether you are drawn to images of families with

minivans or bachelors driving expensive sports cars – whether you react most to logical arguments or emotional messaging.

Assuming these generative methods produce more clicks for advertiser, they will likely become standard practice, and an arms race will follow, with platforms competing to produce the most effective personalization techniques. Even platforms that choose to avoid manipulative tactics for ethical reasons may be pressured to pursue these methods by rivals that take a more aggressive posture on interactive adaptive ads [27]. For this reason, policy solutions may be the only way to prevent the industry from racing in this direction. Furthermore, this risk is likely to accelerate as the market transitions from traditional forms of print and video advertising to new forms of conversational influence.

**Targeted Conversational Influence** can be defined as real-time generative techniques in which pre-defined influence objectives are conveyed to individual users through interactive dialog rather than traditional documents or videos. Until recently, the risks of conversational influence were largely theoretical, as the tools did not exist to enable convincing AI-generated conversations at scale. This has rapidly changed over the last 12 months with the deployment of foundational Large Language Models (LLMs) such as GPT-4 from Open AI.

Currently, the most popular conversational AI systems are text-based chatbots. That said, voice-based systems (powered by LLMs) will become equally pervasive in the near future. We can therefore predict that consumers will soon engage chatbots and voice-bots frequently throughout their daily lives, as third-party developers use APIs to integrate LLMs into apps and websites. For example, in the near future you might request the latest scores on a sports website by speaking naturally to an AI-powered conversational agent. During the interaction, you may receive the desired information (as dialog) but could also be targeted with conversational influence – subtle messaging integrated into the flowing dialog with pre-defined promotional objectives.

In this way, Targeted Conversational Influence may become an insidious form of persuasion, as it may not be easily identified by users as promotional material. Instead, it could be woven skillfully into the dialog, mixed with informational content. And like Targeted Generative Ads, these conversational experiences will be generated in real-time for individually targeted users based on (i) influence objectives provided by sponsors along with (ii) personal background data accessed for the specific user engaged in the dialog. The personal data may include the user's age, gender, interests, hobbies, values, etc... thereby enabling the generative AI system to produce dialog that could optimally appeal to the specific person.

Why conversational influence? As most salespeople know, the best way to persuade a target customer is not to hand them a brochure or tell them to watch a video. Instead, the best way to influence a customer is to engage them in interactive face-to-face conversation so the salesperson can pitch a product, hear the customer's reactions and reservations in real-time, and then adjust the promotional arguments and sales tactics to overcome resistance. It is a cyclic process of pitching and adjusting, all highly individualized, until the salesperson finally talks the customer into the purchase [16, 17, 18].

While conversational pitching has been a purely human skill in the past, generative AI has enabled automated systems to perform all of the steps – (pitching, observing, adjusting) – but potentially with deeper knowledge and greater skills than human representatives. This is dangerous as it could easily create a highly asymmetric relationship between the customer and the AI-powered representative in which the human is deeply outmatched and easily manipulated [16, 17, 22-26]

For example, unlike a human salesperson who only has one persona for engaging customers, a conversational AI is a *digital chameleon* that can assume any speaking style from formal to folksy, nerdy to hip, and can pursue any sales tactic from befriending the customer to playing on their fear of missing out. And the conversational systems will be armed with personal data about every unique user it faces, enabling it to mention the right sports teams or musical groups to ease the user into friendly dialog, or appeal to the right political values or sensibilities.

Furthermore, it is not just about choosing an effective style or tactic for a given user. These personalized conversations will be modified in real time based on user interactions. This means the AI agent could assess the user's reactions, determining which arguments are working and which are being met with resistance. The AI agent will then adjust its arguments and persuasive tactics to maximize appeal and overcome objections. And because tech platforms could document the effectiveness of prior promotional conversations, the system could learn over time which types of arguments work best on specific users. Does that user respond best to logical or emotional appeals? Is that user most influenced by a confident and forceful representative, or a friendly and empathetic one? In our rapidly approaching conversational future, platforms could quickly learn to pull our strings [16].

Of course, the biggest danger is not the optimized ability to sell products and services. The very same techniques could be used to drive propaganda and misinformation, talking users into false beliefs or extreme ideas that they might normally reject. A conversational agent, for example, could easily be directed to convince you that perfectly safe medicines are dangerous. And because it has access to all the world's information, it could cherry-pick evidence in ways that would overwhelm a human user who does not have infinite information access. And even if large LLM providers like Open AI put protections in place, these systems are accessible by API which means that third party developers can implement mediation layers that could easily enable conversations on false or misleading content.

It is therefore critical that policymakers consider the unique dangers of conversational AI and its potential use as a highly personalized and interactive form of targeted influence. It is also important to point out that conversational agents are likely to advance significantly in the next few years, evolving from simple text and voice chatbots to "digital personas" that are embodied as simulated human faces that have an authentic appearance and can express realistic emotional sentiments. These are often referred to as virtual humans or Virtual Spokespeople (VSPs) and are likely to become the interactive face of businesses and services that are accessed online [17]. They are also likely to be more impactful and potentially more manipulative than current chatbots because they will be perceived as more human and will convey information with greater affect [19,20, 24].

Some argue that interactive and adaptive influence on an individual basis is not a new problem because human salespeople and other representatives perform similar tactics. There are two counterarguments. First, while human persuaders can be quite effective, they cannot be deployed at scale with nearly the flexibility or diversity. Second, and more importantly, when human representatives engage in real-time dialog with a human target, the power balance is mostly reciprocal, with each party having similar ability to read the other and similar access to supporting facts. When a human is engaged with an AI-powered conversational agent, the power balance has the potential to be extremely asymmetrical. This is especially true if the AI agent is *designed for persuasion* (i.e., trained on sales tactics, personality types, cognitive biases, and other influence related skills). This is discussed in detail in the following section.

### 3. THE ASYMMETRY OF CONVERSATIONAL AI

When a user is engaged in real-time dialog with an AI-powered conversational agent, the power balance has the potential to be highly skewed with the human at significant disadvantage. These imbalances can be summarized as follows:

**Familiarity Asymmetry:** unless regulated, it is likely that AI agents will have access to personal data about target users, ranging from age, interests, education and political views to their favorite teams, movies, and musical artists. This will empower the AI to customize dialog for individual targets. On the other side of the conversation, the human will know nothing about the entity they are conversing with. And if the AI agent is given a visual or vocal persona that represents a particular age, gender, style, or background – it is entirely a façade. In this way, the AI agent could know a great deal about the user, while the user would know nothing about the AI agent. This is a deeply asymmetric relationship and a scenario that has no equivalence with respect to human salespeople.

**Emotional Asymmetry:** unless regulated, it is likely that AI agents will be able to “read users” when engaged in conversation, assessing not just resistance or agreeableness to influence, but also sense emotional state. *Is the user getting angry, anxious, excited?* This is detectable in text-based dialog and will become even more invasive in voice-based systems in which AI agents can assess vocal inflections of target users. The human, on the other hand, will be unable “to read” their digital counterpart, for any emotions or inflections conveyed in language, voice, or facial expressions (in the case of virtual personas) are entirely fictional, chosen to maximize impact and do not reflect emotional state. This is a deeply asymmetric relationship and not something that has an equivalent with respect to human salespeople.

**Continuity Asymmetry:** unless regulated, platforms deploying generative media will likely keep track of user’s reactions to targeted influence and will learn what types of persuasive tactics are most effective on individual users. The human in the loop, on the other hand, will learn nothing about the digital representatives they engage throughout their day, for they could be *digital shapeshifters* that can assume any style, tactics, or persona in response to simple API calls. This too is a deeply asymmetric scenario that has no equivalent with human salespeople.

**Information Asymmetry:** unlike human influencers who can make arguments and counterarguments based on human-level knowledge and experience, AI agents will be able to craft dialog that draws on a nearly infinite information pool and could easily cherry-pick points that the human could not possibly validate in real-time. In fact, an AI agent could create the illusion of expertise by citing overwhelming factual information as a deliberate form of persuasion. This is a deeply asymmetric situation that has no direct equivalent with human salespeople.

**Strategic Asymmetry:** unless regulated, conversational agents could be trained in sales tactics, negotiation tactics, human psychology, cognitive biases, game theory and other persuasive strategies that make them extremely effective instruments of influence. For example, in 2022 DeepMind developed a strategic AI system called DeepNash that learned to bluff human players and sacrifice gaming pieces for the sake of long-term victory [21]. While human salespeople are often trained in sales tactics, it is likely that AI agents will possess a broader and more nuanced range of persuasive strategies than any human could command

creating a uniquely asymmetric power-imbalance.

Clearly, generative media poses unique interactive risks that have not been faced in traditional media. And as the digital world transitions to conversational interaction with apps, websites, and other software tools, the dangers will only increase. To help policymakers appreciate the unique risks posed by interactive and adaptive forms of influence, it is useful to describe the issue in the language of “feedback control,” which is a common engineering technique used to drive real-time behaviors of target systems. As will be described in the section below, many dangers associated with generative media relate to a platform’s ability to “close the loop” around individual users and use feedback control to maximize influence. Appreciating the basics of control theory may help guide effective policy protections.

### 4. THE AI MANIPULATION PROBLEM

The AI Manipulation Problem refers to real-time scenarios in which artificial agents manipulate human users by imparting targeted influence through persuasive interactions [2, 3, 17]. It abstracts the problem down to the following sequence of steps performed by an AI system on an individual human user:

- 1) Impart real-time influence on an individual user;
- 2) Detect the user’s reaction to the imparted influence;
- 3) Adjust influence tactics to increase persuasive impact;
- 4) Repeat steps 1-3 to maximize user influence over time;

This may sound like an abstract series of computational steps, but we humans often just call it a conversation. As described above, when a human wants to influence another human in real-time, they will make arguments (i.e., impart targeted influence), listen to reservations and counterarguments (i.e., sense user reactions) and then adjust tactics to overcome objections. The unique danger is that generative AI systems can now perform all of these steps, creating a “feedback control system” around a human user that has the potential to be an extremely effective form of coercion, manipulation, and persuasion.

To help policymakers appreciate how a “feedback control system” is fundamentally different from traditional forms of media-based influence, it is helpful to provide background from the field of Control Theory (see Figure 1).

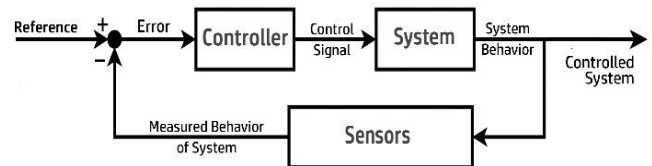


Figure 1: Standard Diagram of Feedback Control System

Control Theory (CT) is the engineering discipline that formalizes how real-time feedback loops can be used to guide the behavior of any system towards a specified goal. The classic example is a thermostat. A homeowner sets a temperature goal and if the house falls below the goal, a controller turns up the heat. If the temperature rises too much, the controller reduces the heat. When working properly, the thermostat keeps the house close to the specified objective.

Referring to Figure 1, the System being controlled in the heating example is a house, the Sensor is a thermometer, and the Controller is a thermostat that modulates the heat as needed. An input signal called the Reference is the temperature goal. The goal is compared to the actual temperature in the house (i.e., the

Measured Behavior) and the difference is fed into the controller which then determines how to adjust the heat. This creates a real-time feedback loop that continually detects behavior (i.e., temperature) and imparts influence (i.e., heat).

While a controller can be as simple as a thermostat, it can also be quite sophisticated. For example, self-driving cars use AI-controllers to navigate traffic, achieving targeted goals in rapidly changing environments. When considering the use of generative AI to impart influence on humans, similar feedback-control methods can be employed. Instead of a simple thermostat that turns up or down the heat as needed, an AI system could modulate persuasive tactics used on an individual human based on the Measured Behavior of the system, which in this case are the real-time reactions of that user. The Reference signal, instead of being a temperature goal, becomes the influence objectives of a third party such as a corporate sponsor or state actor. The third party need not be the creator of the generative AI system but merely needs to input influence objectives through API access. This yields the manipulative control system of Figure 2 below.

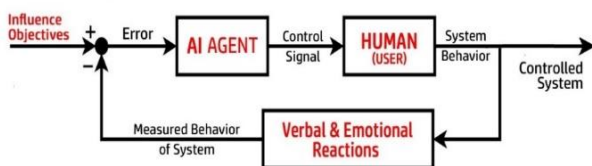


Figure 2: Manipulative Control System using Generative Media

As diagramed above, AI-driven control systems can easily be deployed using generative media techniques to create interactive agents that are designed to maximize the persuasive impact on individually targeted users. For example, conversational agents could be designed to draw target users into friendly dialog and gradually *talk them into* specific influence objectives. This is instructional for policymakers because restricting one or more of the elements in the control system diagram above, the feedback loops could be disrupted and the ability of third parties to deploy optimized influence could be mitigated.

For example, regulators could consider banning generative AI systems from making any real-time adjustments to promotional imagery or promotional messaging in response to detected user reactions. This could prevent generative AI systems from being used in these highly manipulative ways.

In addition, regulators and policymakers could require that all forms of Targeted Conversational Influence be obligated to disclose their *influence objectives* to the user upon initiation of the dialog. This disclosure would ideally be part of the conversation and not a subtle disclaimer that can be skipped over. This approach could prevent conversational influence from being added covertly to conversational exchanges without the user knowing that a third party has directed the system to impart personalized influence. This is significant because it will allow the user to view the interaction as promotional and bring a healthy dose of skepticism to the conversational exchange.

And finally, policymakers and regulators should consider an outright ban on all forms of Interactive Generative Influence that targets minors, for children are more impressionable than adults and already have a significantly harder time distinguishing between promotional and authentic content. This is especially true for conversational influence targeting minors.

## 5. POLICY RECOMMENDATIONS

As described above, many new risks emerge when targeted influence becomes interactive, adaptive, intelligent and highly

personalized. To mitigate those risks, three categories of policy recommendations are discussed below under the headings of Interactive Transparency, Emotional Privacy, and Protecting Human Agency. These are not the only policy needs related to generative AI but are directed to the unique and new risks that emerge from Interactive Generative Media.

### 5.1 Interactive Transparency

Marketing is pervasive in our physical and digital worlds, but most adults can easily identify promotional content across a wide range of environments. This allows individuals to consider ads in the appropriate context – as paid messaging delivered by a third-party sponsor. This enables consumers to bring skepticism and critical thinking when considering messaging for products, services, political ideas, and other promotional content. When engaging with conversational interfaces, from current chatbots and voice-bots to upcoming video-based “digital humans,” marketers and propagandists could easily undermine our ability to contextualize promotional material by easing it into conversations, blurring the boundaries between informational dialog and promotional messaging injected on behalf of a paying sponsor or state actor.

**Regulatory Approach:** To adequately safeguard the general public, regulators and policymakers should consider requiring all conversational systems (text, voice, or video) to disclose any influence objectives that may have been assigned by a third party. This disclosure should be expressed by the conversational agent, informing the user that it is transitioning from informational content to promotional messaging and indicating what its influence objective is. For example, when providing factual information about the location of EV charging stations, an AI-powered chatbot should be required to inform the user if it transitions the conversation to promotional content about the benefits of a particular sponsored vehicle. With such disclosure, consumers will be less likely to confuse informational content and promotional messaging within a single stream of dialog. This is especially important for children who already struggle to identify promotional content across various settings.

### 5.2 Emotional Privacy

We humans communicate by expressing emotions on our faces, in our voices, and through postures and gestures. Emerging AI systems can now identify human emotions from human faces, voices, and bodies as well as from eye motions, pupil dilations and other physiological responses. While emotion detection will enable computers to communicate more effectively with humans when engaged in voice and video chat, it has significant risks. That’s because AI systems can detect emotions that are not perceptible to human observers. For example, AI systems can detect “micro-expressions” that are too brief or subtle for humans to notice. Even more concerning, AI agents can be designed to adjust promotional tactics mid-dialog based on detected emotions. These cues, whether detected in facial or vocal features, or inferred from the language expressed by users, can be used to determine which persuasive tactics are eliciting the most positive reactions and which are being met with resistance. This can allow the generative AI system to adapt its tactics in real-time for optimized influence.

**Regulatory Approach:** In the context of interactive systems, regulators should ban AI-powered emotional assessments that exceed natural human observational abilities. This would mean

not allowing vital signs, pupil dilation, micro-expressions, or facial blood patterns to be used in emotion detection by AI systems. In addition, the risk to users is greatly amplified if platforms are allowed to collect emotional data over time and create emotional profile models. Emotional models of individual users would allow AI systems to predict user reactions to a wide range of stimuli. This could be used in highly manipulatory ways. And finally, regulators should restrict or ban the use of real-time emotional assessments by interactive conversational systems that have a promotional agenda or other targeted influence objectives.

### 5.3 Protecting Human Agency

**Background:** As described above, real-time generative AI can be used for cognitive and/or behavioral manipulation on human users through real-time feedback control. This can be achieved by conversational AI systems that perform the following steps: (i) impart targeted influence on an individual user, (ii) detect the user's emotional or behavioral reaction to that influence, and (iii) repeatedly adjust the influence and detect results to gradually maximize the persuasive impact. This could easily cross the line from marketing to manipulation and compromise human agency.

**Regulatory Approach:** Regulators and policymakers should consider aggressive regulations that restrict or ban the use of AI-powered feedback-loops that adapt real-time content delivered to individual in response to their behavioral or emotional reactions. This is especially true for conversational content deployed through interactive dialog but could also involve adaptive forms of image-based and video-based advertising. Such protections are critical to guard against cognitive, emotional, or behavioral manipulations of individually targeted users at scale.

## 6. CONCLUSIONS

For the reasons described above, policymakers and regulators must expand their risk assessment of generative AI beyond its ability to create traditional misinformation at scale. In particular, policymakers must also consider the unique risk of generative AI being used to unleash a new form of media that is personalized, interactive, and adaptive in real-time. Referred to herein as Interactive Generative Media, this new form of targeted influence is particularly dangerous because it can employ feedback control methods to optimize persuasive impact on individual users. If regulatory protections are not put in place that focus specifically on AI-powered interactive and adaptive media, it is likely that two dangerous techniques will become widely used by large technology platforms: “*Targeted Generative Advertising*” and “*Targeted Conversational Influence*.” The potential harms could be mitigated if policymakers adopt meaningful guardrails that prevent real-time AI-powered systems from “closing the loop” around human users for persuasive purposes.

## 7. REFERENCES

- [1] Giovanni Spitale et al., AI model GPT-3 (dis)informs us better than humans. *Sci. Adv.* 9, eadh1850 (2023). DOI:10.1126/sciadv.adh1850
- [2] Rosenberg, Louis. “The Metaverse and Conversational AI as a Threat Vector for Targeted Influence.” 2023 IEEE 13th Annual Computing and Communication Workshop and Conf. (CCWC), 2023. [10.1109/ccwc57344.2023.10099167](https://doi.org/10.1109/ccwc57344.2023.10099167).
- [3] Rosenberg, Louis. The Manipulation Problem: Conversational AI as a Threat to Epistemic Agency. 2023 CHI Workshop on Generative AI and HCI (GenAICHI 2023). Association for Computing Machinery, Hamburg Germany (April 28, 2023) [link](#)
- [4] Gunn, H., Lynch, M.P.: The internet and epistemic agency. In: Lackey, J. (ed.) *Applied epistemology*, pp. 389–409. Oxford University Press, Oxford (2021).
- [5] Bozdog, Engin & van den hoven, Jeroen. (2015). Breaking the filter bubble: democracy and design. *Ethics and Information Technology*. 17. 10.1007/s10676-015-9380-y.
- [6] Caplan, B. (2008). *The myth of the rational voter: Why democracies choose bad policies*. New edition. Princeton, NJ; Woodstock: Princeton University Press. <http://www.amazon.com/The-Myth-Rational-Voter-Democracies/dp/0691138737>
- [7] Hardin, R. (2009). *Deliberative Democracy*. In T. Christiano & J. Christman (Eds.), *Contemporary debates in political philosophy*. West-Sussex: Blackwell.
- [8] Coeckelbergh, M. *Democracy, epistemic agency, and AI: political epistemology in times of AI*. *AI Ethics* (2022). <https://doi.org/10.1007/s43681-022-00239-4>
- [9] Commission on Information Disorder Final Report, Nov 2021. Aspen Institute.
- [10] Jae Kook Lee, Jihyang Choi, Cheonsoo Kim, Yonghwan Kim, Social Media, Network Heterogeneity, and Opinion Polarization, *Journal of Communication*, Volume 64, Issue 4, August 2014, Pages 702–722, <https://doi.org/10.1111/jcom.12077>
- [11] Tucker, J.A., et. al, 2018. Social media, political polarization, and political disinformation: A review of the scientific literature. *Political polarization, and political disinformation: a review of the scientific literature* (March 19, 2018).
- [12] Rosenberg, Louis. (2022). *Social Media is Making Us Stupid, but We Can Fix It*. VentureBeat. Retrieved January 1, 2023, from <https://venturebeat.com/business/social-media-is-making-us-stupid-but-we-can-fix-it/>
- [13] Murphy, Hannah, and Cristina Criddle. “Google to Deploy Generative AI to Create Sophisticated Ad Campaigns.” *Financial Times*, April 20, 2023.
- [14] Mehta, Ivan. “Meta Wants to Use Generative AI to Create Ads.” *TechCrunch*, April 5, 2023.
- [15] Person, and Sheila Dang. “Exclusive: Microsoft's Bing Plans AI Ads in Early Pitch to Advertisers.” *Reuters*. Thomson Reuters, February 17, 2023.
- [16] Rosenberg, Louis (2023). “Conversational AI Will Learn to Push Your Buttons.” *Barron's*. February 23, 2023.
- [17] Rosenberg, L., 2022, October. Marketing in the Metaverse and the Need for Consumer Protections. In 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0035-0039). IEEE. [10.1109/UEMCON54665.2022.9965661](https://doi.org/10.1109/UEMCON54665.2022.9965661)
- [18] Rosenberg, Louis (2023). “The Profound Danger of Conversational AI.” *VentureBeat*, February 4, 2023.
- [19] Breves, Priska. “Biased by Being There: The Persuasive Impact of Spatial Presence on Cognitive Processing.” *Computers in Human Behavior*, vol. 119, 2021, p. 106723., <https://doi.org/10.1016/j.chb.2021.106723>
- [20] Han, E., Miller, M.R., DeVeaux, C., Jun, H., Nowak, K.L., Hancock, J.T., Ram, N., Bailenson, J.N. (December, 2022). People, Places, and Time: A Large-scale, Longitudinal Study of Transformed Avatars and Environmental Context in Group Interaction in the Metaverse. *Journal of Computer-Mediated Communication*.
- [21] Perolat J, et. al., Mastering the game of Stratego with model-free multiagent reinforcement learning. *Science*. 2022 Dec 2;378(6623):990-996. doi: 10.1126/science.add4679.

- [22] Rosenberg, Louis. "Why Generative AI Is More Dangerous than You Think." VentureBeat, May 6, 2023. Available at: <https://venturebeat.com/ai/why-generative-ai-is-more-dangerous-than-you-think/>
- [23] Robertson, Derek. (2022, September 14). "The Most Dangerous Tool of Persuasion." POLITICO. POLITICO. <https://www.politico.com/newsletters/digital-future-daily/2022/09/14/metaverse-most-dangerous-tool-persuasion-00056681>
- [24] Heller, Brittan and Bar-Zeev, Avi. "The Problems with Immersive Advertising: In AR/VR, Nobody Knows You Are an Ad", Journal of Online Trust and Safety, October 2021.
- [25] Rosenberg, Louis. "Mind Control: The Metaverse May Be the Ultimate Tool of Persuasion." VentureBeat, VentureBeat, 22 Oct. 2022, <https://venturebeat.com/virtual/mind-control-the-metaverse-may-be-the-ultimate-tool-of-persuasion/>
- [26] Heller, Brittan. "Watching androids dream of electric sheep: immersive technology, biometric Psychography, and the law." Vand. J. Ent. & Tech. L. 23 (2020)
- [27] Spisak, Brian., Rosenberg, Louis and Beilby, Max (2023) 13 Principles for Using AI responsibly, Harvard Business Review. June 30, 2023. <https://hbr.org/2023/06/13-principles-for-using-ai-responsibly>.