

Introduction

A Convergence Approach to the Mis/Disinformation Problem

DAVID CASTILLO AND SIWEI LYU

The unprecedented spread of false, fake, and misleading information in our media ecosystem is the flip side of the internet's promise of universal access and information democratization. As an unwanted consequence of the neck-breaking advances of digital technology and artificial intelligence, and the emergence of influential social media platforms, we are experiencing online mis/disinformation with escalating speed, volume, and level of sophistication. The omnipresence of misinformation in our environment is eroding our trust in legitimate sources of information and posing significant threats to consumers/users everywhere in the world. To fully understand the challenges facing our communities and to explore effective strategies for individual and collective action, we need the long view and broad perspective that only collaborative and multidisciplinary approaches (including the full range of scientific, technical, and humanistic fields) can bring.

For context, it is important to keep in mind that mis/disinformation has accompanied the rise of new media throughout history as an unfortunate byproduct of the modern world. The spread of printed materials in the 1500s provided the expanding reading public access to a vast repository of information, but it also made way for opportunists, propagandists, and demagogues, who would quickly learn to exploit the vulnerabilities of the expanding print market to their advantage. Early Modern Europe was flooded with sensationalist and scandalous tales of supernatural occurrences,

phony or fake news stories, and baseless conspiracy theories. The “yellow journalism” of the late 1800s, and the arrival of radio and audiovisual media in the first half of the twentieth century are also notable examples of the correlation between the expansion of mass media and the spread of mis/disinformation, including the kind of orchestrated propaganda we associate with the Info Ops of the brutal totalitarian regimes of the 1900s. Yet, the speed of transmission and the volume of mis/disinformation we see in the twenty-first century are without parallel in human history. The exponential increase of corrupt, misleading, manipulative, and exploitative information that’s flooding our media-saturated world is causing significant harm to individuals and communities, eroding our trust in democratic institutions, and weakening our ability to respond to emerging crises, such as the recent pandemic or the looming threats of climate change.

The current emergency is at least in part a consequence of the ruthless efficiency of the powerful algorithms that media giants use to optimize user engagement and maximize profits. The data-gathering artificial intelligence (AI) tools employed by Facebook, Twitter (now X), TikTok, and other social media companies work behind the scenes to amplify and personalize sensationalist and polarizing content that attracts more views. Whatever gets a response spreads, and bad actors exploit this vulnerability. It is important to understand that the digital media landscape today is far different from that of the late 1990s, when internet users accessed content through search engines such as Lycos and web directories such as Yahoo. Those early internet services had no embedded mechanism to promote content. Consuming content required users to intentionally search for a keyword or browse to a particular website or forum. That era is basically gone. Instead, today’s social media feeds serve up fringe and polarizing content, which users may not be intentionally searching for, thanks to the tech industry’s embrace of two key technological developments: personalization, spurred by mass collection of user data through web cookies and big data systems, and algorithmic amplification, the use of powerful AI instruments to select the content that will be shared with individual users based on their interests and search history.

What this means in practical terms is that we can effectively live inside our own media silos, a reality of our choosing that’s confirmed on an hourly basis by our AI-curated feeds. In the most extreme cases, these alternative realities can become rabbit holes populated by self-sustaining illusions. One particularly troubling type of audiovisual illusion capable of spreading mis/disinformation and locking us inside AI-generated versions of the world is

the DeepFake, featured with increased frequency in recent news and social media posts. The fabrication and/or manipulation of digital media is not a new phenomenon. The synthesizing of realistic audio, images, and videos using algorithms has been an essential task in such fields as signal processing, computer graphics, and computer vision for some time. Yet, with pre-AI tools the creation process was lengthy, costly, and technically demanding for ordinary users. In recent years, however, the developments of AI technologies have significantly lowered the requirements on resources, time, and technical expertise for the creation of convincing fakes. The unprecedented computing power and the powerful AI technology available today, deep neural networks (DNNs) in particular, have made it easier, cheaper, and much less time consuming to generate sophisticated and compelling fakes.

DeepFakes first caught the public's attention in late 2017 when a Reddit account called DeepFake, a portmanteau of deep learning and fake media, started to spread pornographic videos with transplanted celebrity faces created using a DNN-based algorithm. Since then, more sophisticated algorithms to synthesize realistic audio, images, and videos have emerged, along with many open-source software tools and commercial services. In essence, DeepFakes are the tip of the iceberg of the troubling trend of increasingly realistic online mis/disinformation. By creating false records of the words and actions of individuals, DeepFakes can cause significant harm when weaponized. For instance, a fake video showing a politician engaged in inappropriate activity could be enough to sway an election if released at the right time. A falsified audio recording of a high-level executive commenting on her company's financial situation could send the stock market awry. Using a synthesized realistic human face as the profile photo for a fake social platform account can significantly increase the effectiveness of deception schemes. An online predator can masquerade as a family member or friend in a video chat to lure unsuspecting victims.

Left unchecked, DeepFakes can escalate the threat of online disinformation and fundamentally erode our trust in digital as well as traditional media, including reputable news sources and legitimate educational materials. In addition, DeepFakes pose a considerable threat to our cognitive security and could be used in combination with other cyberattacks to breach cyber systems. Working as a form of manipulative disinformation, DeepFakes can "hack" our perceptual system and decision-making process and endanger our personal data. Impersonating someone has become much easier with audiovisual synthesis. We have already seen a few cases of DeepFakes with devastating real-world consequences. GAN-generated face images have been

used as profile photos for fake accounts on social platforms such as Twitter, Facebook, Instagram, and LinkedIn. In 2020 alone, 4,000 fake accounts were found on these social media platforms. Using such realistic face images as profile photos significantly increases the deceptiveness of those fake accounts. Another incident involved a scammer who successfully used a synthesized voice to impersonate the CEO of a UK company and mislead an employee to wire transfer a substantial amount of money to the scammer's bank account. Reports show that hackers have used DeepFakes to falsify biometric data to gain access to essential information systems. A particularly insidious case was a DeepFake video of the Ukrainian President directing the Ukrainian troops to surrender to the Russians. The video circulated on social media and Ukrainian news websites before being debunked and removed.

The security threats posed by the availability of increasingly effective DeepFake technologies have received broad attention from lawmakers and government officials worldwide. In the United States, Congress has passed several bills to regulate the use of DeepFakes, including the Malicious Deep Fake Prohibition Act of 2019, the DEEP FAKES Accountability Act, the Deepfake Report Act, and the IOGAN Act, 2019. The European Union and China have engaged in similar legislative efforts. Major social media platforms (e.g., Twitter and TikTok) have followed suit to generate policies aimed at controlling the production and spread of DeepFakes.

As serious as the danger of manipulative DeepFakes truly is, this is by no means the only threat posed by the groundbreaking advances we are witnessing in the field of artificial intelligence. The debates surrounding the recent release of ChatGPT seem to suggest a worsening of the problem in the short term due to the potential of AI technologies to generate convincing illusions that could lock unsuspecting users into alternative versions of reality. As Stuart A. Thompson, Tiffany Hsu, and Steven Lee Myers wrote in a *New York Times* article published on March 22, 2023, “Even as tech giants scramble to join the commercial boom prompted by the release of ChatGPT, they face an alarmed debate over the use—and potential abuse—of artificial intelligence. The technology’s ability to create content that hews to predetermined ideological points of view, or presses disinformation, highlights a danger that some tech executives have begun to acknowledge: that an informational cacophony could emerge from competing chatbots with different versions of reality, undermining the viability of artificial intelligence as a tool in everyday life and further eroding trust in society” (Thompson, Hsu, and Myers 2023).

This is where the example of Miguel de Cervantes's most famous character comes in handy for us as a sort of literary analogue and critical caricature. Don Quixote is of course known for living inside a self-sustaining fantasy world (his own chosen reality informed by chivalric illusions) inside of which ordinary windmills can become evil giants. His fictional exploits are the inspiration behind the expression "tilting at windmills," meaning "attacking imaginary enemies or evils." If we can make the leap from the pages of Cervantes's famous novel (and the circumstances of the cultural crisis that inspired it) back to our troubled present, we can say that our media environment is breeding an alarming number of Quixote-style vigilantes (let's call them Q-Knights) intent on charging at all manner of things based on the self-sustaining illusions of conspiracy theories they read about in their media feeds. Like Don Quixote himself, the Q-Knights of our time can live inside a self-sustaining illusion in which the proliferation of mattress stores in close proximity to each other is clear evidence of a money-laundering operation or other nefarious activity possibly linked to a government agency (Hanbury 2018) and the basement of a popular pizza restaurant in Washington, DC, is the headquarters of a child-trafficking ring (BBC Trending 2016). They know what's really cooking "down there," even if there's no "down there" there, because they've read about it online. The fact that the restaurant in question has no basement is only proof of the vast powers of the ring leaders and their accomplices, who managed to cover their tracks so effectively. The resilience of these conspiratorial illusions even when challenged by seemingly incontrovertible evidence is indeed reminiscent of Don Quixote's trademark response to the crashing reality of windmills in the famous scene of the 1605 volume:

"God save me!" said Sancho. "Didn't I tell your grace to watch what you were doing, that these were nothing but windmills, and only somebody whose head was full of them would think otherwise?"

"Be quiet, Sancho . . . replied Don Quixote. "Matters of war, more than any others, are subject to continual change; moreover, *I think, and therefore it is true*, that the evil enchanter Frestón . . . has turned these giants into windmills in order to deprive me of the glory of their defeat: such is the enmity he feels against me; but in the end, his evil arts will not prevail against the power of my virtuous sword." (146, our translation, our emphasis)

An undeterred Don Quixote would go on to wage war against all manner of evil powers, which he “recognizes” in things that look like wineskins, watermills, and even livestock. In chapter 21 of the novel, Don Quixote and Sancho come across a human figure who is wearing something shiny on his head. The narrator identifies the man as a barber on a mule wearing a barber’s basin as a makeshift hat to protect himself from the pouring rain, but Don Quixote begs to differ. The approaching horseman cannot be but a rival knight who is obviously in possession of Mambrino’s magical helmet, which Don Quixote is destined to win in battle. When the unsuspecting barber sees Don Quixote charging at him full tilt, he dismounts in a hurry and runs for his life, leaving behind his ride and his headpiece. This is what happens next:

When Sancho heard the barber’s basin being called a helmet, he could not contain his laughter . . .

“What are you laughing about, Sancho?” said Don Quixote.

“I’m laughing,” responded Sancho, “thinking about the big head belonging to the pagan owner of this helmet, which looks perfectly like a barber’s basin.”

“Do you know what I think, Sancho? I think this famous fragment of this enchanted helmet, by some strange accident must have come into the hands of someone who could not recognize or appreciate its value, and without realizing what he was doing, seeing that it was cast from the purest gold, must have melted the other half for its worth, and from this half he made what looks like a barber’s basin, just as you say. *But no matter, since I know what it is, its transformation makes no difference.*” (260, our translation, our emphasis)

It bears repeating, for one who “knows” what things really are regardless of their appearance, no amount of fact-checking or “debunking” by the Sanchos of the world will mean a thing. As for the familiar Q-knights of our own day who may consider attacking pizza parlors and mattress stores, they likewise know “the truth” of what is really going on, and no amount of evidence to the contrary will dissuade them otherwise. This is why efforts to combat mis/disinformation and conspiracy theories today and to mitigate their devastating impact in our communities cannot rely on fact-checking and debunking approaches alone. In essence, it is very hard to debunk or fact-check someone out of a conspiratorial rabbit hole, which is why the

better approach might be to prevent people, including ourselves, our family members, our virtual communities from falling into those rabbit holes in the first place. But how?

One thing we have learned in the last few years is that there's no technical magic bullet capable of solving the problem of online deception and manipulation on its own. Domain experts in media technologies are the first to admit that even the most effective technical fixes will be at best partial and temporary since they would not be able to get to the root of the mis/disinformation problem. The available research on DeepFake-driven disinformation is a case in point. The mounting concerns over the nefarious use of DeepFakes have spawned increasing interest in counter technologies, with substantial support from government and private companies. Notable examples include the DARPA MediFor and SemaFor programs, the NIST 2018, 2020, and 2021 Synthetic Data Detection Challenge, and the DeepFake Detection Challenge (Meta AI 2020), sponsored by Facebook, Microsoft, Amazon, and Partnership in AI. DeepFake forensics has become an active research area in the past few years in response to the growing concerns. Indeed, it is fair to say that the current efforts in DeepFake forensics heavily tilt toward detection, often looking at the problem in terms of a simple binary classification: real or fake. Thus, we have more than 300 publications focused on this type of classification relying on more than fifty methods with code and datasets. Yet some studies (including Shane et al. 2021) show that the straightforward approach of labeling DeepFakes as part of debunking operations can actually undermine mitigation efforts by not taking sociopsychological considerations into account—users are drawn by curiosity to watch DeepFakes, which increases attention and promotes the spread of the labeled fake content. This is consistent with psychological research data that describes the vulnerabilities and failures of users when facing disinformation while leaving little space for potential solutions (Khodabakhsh et al. 2018; Lago et al. 2021). Taken together, these studies show that we need to think about best practices in intervention procedures, even after a DeepFake is exposed.

Given the complex nature of the challenge and the limited reach of purely technical fixes, we argue that there is a compelling need to refocus our approach, bringing together experts from computer science and engineering with researchers in social and behavioral sciences and the arts and humanities, along with representatives of the user communities. Convergence research, which is inherently multidisciplinary, problem-focused, and solutions-based, offers an ideal framework for the development of a more coherent and

comprehensive approach to help guide what we study, whom we study, how we conduct our research, and who needs to be involved in the research process itself. These notions inspired the recent creation of the UB Center for Information Integrity. Our more than forty faculty members recognize that tackling what is clearly a defining challenge of our time requires multi-disciplinary teams working together to supplement existing red-flagging and debunking practices with innovative preventative or “pre-bunking” approaches with the goal of raising awareness and increasing resilience. As consumers and users of complex digital products that compete for our attention, we must be able to identify dubious and malicious sources of information and to develop strategies to expose false claims and narratives and build trust in our communities to make effective interventions.

The present collection of essays is an attempt to model a convergence approach involving the Center for Information Integrity and the Humanities Institute at the University at Buffalo. We view this scholarly collaboration as a first step in our efforts to increase awareness and build resilience against the omnipresent threat of mis/disinformation in our media environment. We have structured this volume into three parts dealing with complementary and interrelated issues. Part I, “Misinformation and Artificial Intelligence,” deals most directly with the threat of mis/disinformation in the context of what Shoshana Zuboff has called “surveillance capitalism.” The essays included in this section reflect on the danger of outsourcing judgment and decision making to AI instruments in key areas of public life, from the processing of loan applications to school funding, policing, and sentencing. Part II, “Science Communication, Cultivating Awareness,” is focused on the need to rethink how scientific findings are communicated to the public. We suggest that scientists need to cooperate with colleagues in other disciplines and community representatives to help minimize the negative effects of mis/disinformation in such vital areas as climate change science and public health. The consensus is that no amount of science-explaining on its own will likely work in the absence of a concerted effort to listen to and incorporate community questions, concerns, needs, and aspirations. Finally, part III, “Building Trust,” expands on the issues of the previous sections to advocate for and explore instances of trust-building initiatives as a necessary precondition of community-oriented scholarly activity and effective intervention strategies in high-impact areas such as public health. The volume closes with an Afterword by Christina Milletti, executive director of the UB Humanities Institute, in which she foregrounds the power of fiction to hone critical-thinking skills and build awareness and resilience to misinformation.

In the first essay of the collection, E. Bruce Pitman, an expert in mathematical modeling, explains how our trust in AI systems to make critical decisions can lead to fundamentally erroneous, unjust, and potentially catastrophic outcomes for individuals and communities. As he writes, “AI systems are built by humans, are terribly fallible, and do not ‘learn,’ no matter how generously one defines ‘learn,’ from their experiences . . . AI systems can provide insight that can be valuable to human decision-makers. But we should not have unwavering confidence in AI systems, on their own, to behave predictably and interpretably.” Pitman provides examples of critical areas in which decision making should never be left to unchecked or unsupervised AIs. These areas include the criminal justice system, the financial system, the medical system, and the social services system. He argues that “interpretability of the models used to assist in these critical sectors should be a bare-minimum requirement.”

Comparative literature professor Ewa Ziarek takes the baton in chapter 2 with a broad-ranging discussion of disinformation, power, and the automation of judgments with a focus on “harms to democracy.” Her starting point is that the current disinformation emergency cannot be understood in isolation. Instead, our disinformation problem must be examined in the larger context of ongoing threats to democracy posed by the rise of new technologies of power emerging from the conjunction of big data, digital capitalism, and the outsourcing of political judgments to algorithmic procedures. In this sense, the current mis/disinformation emergency is but a symptom of the momentous computational transformation of our societies in which algorithms play increasingly critical roles in virtually all areas of human activity, from communications and entertainment to information gathering and ranking, to hiring, banking, health care, and dating.

Ziarek coincides with Pitman in foregrounding the lack of transparency and interpretability of algorithmic decision making, a consequence of which is the erosion of democratic processes inside increasingly opaque societies: “Because algorithmically driven global practices of data collection, user profiling, surveillance, and predictive analytics operated by the digital technology giants—Amazon, Google, Facebook, Microsoft, Apple—are not open to public scrutiny, they constitute ‘black box societies’ (Pasquale 2015), in which disinformation, distrust, and conspiracies spread ‘like a virus.’ This global hegemony of private digital technology corporations not only raises questions of data privacy and increased surveillance, but also risks transforming politics and demands for justice into what other scholars and philosophers have called ‘algorithmic governmentality’ (Stiegler 2019; Rouvroy and Barnes, 2013), in which political and juridical decisions are

increasingly replaced by automated algorithms.” The risks here are numerous, not the least of which is the replacement of deliberative processes by opaque practices of digital profiling and decision making. The virtual impenetrability of algorithmic modeling makes it harder to question assumptions, challenge potential biases, and impugn faulty processes and unjust outcomes. We would argue that this is indeed the larger context within which the “integrity” of information, its communicability and applicability, should be pondered as we travel deeper into an AI-assisted future.

In chapter 3, Yotam Ophir, Raphaela Velho, and Lilian Tzivian reflect on the politicization of science and lack of compliance with science-consistent recommendations in such critical areas as public health, with the recent pandemic as a glaring illustration. They discuss and ultimately reject the Deficit Model, which explains inadequate compliance with science as the result of a lack of understanding. Instead, they rely on available data from several countries and regions of the world to demonstrate that it wasn’t ignorance but politically motivated distrust that most “hampered the global effort to slow down the COVID-19 epidemic.” This is why, they argue, scientists must work to (re)build trust, not just by improving their communication strategies and being more transparent about the nature of scientific research but, most importantly, by paying attention to the needs and motivations of their audiences. This requires a better understanding of motivated reasoning and a more inclusive view of potential communication partners: “Importantly, understanding the role played by motivated reasoning in the rejection of science in the face of scientific facts can open doors to more effective science communication (Bisgaard 2019). Groups such as the Evangelical Climate Initiative (Nagle 2008), whose messages explain how protecting the Earth and fighting climate change could be understood through Christian values and perspective, provide an example for effective science communication targeting value-motivated groups and individuals.”

In chapter 4, filmmaker and podcaster John Fiege advocates for “alternative” approaches to science communication in the face of resistance to climate action by politicians, business owners, media personalities, and the public at large. His own approach, which he effectively rehearses in his podcast and films, may be best described as deeply personal, inclusive, and conversational: “This essay is about what I have learned in my search for an alternative way to communicate and make media about the ecological crisis—one that brings people together and engages our best thinking, rather than pushing us apart. I am a filmmaker focused on confronting

environmental problems through the stories of communities on the front lines of environmental calamity. Feeling the urgency to increase the number of stories I can address, the speed of production, and the frequency of my engagement with audiences, I recently launched a podcast, called *Chrysalis*, as a way to confront our ecological predicament through conversations with a broad diversity of environmental thinkers. As the title intimates, *Chrysalis* is a podcast about transformation—the transformation of the podcast guests, the audience, our collective relationship to the rest of nature, and me personally as I engage deeply in conversation.”

In chapter 5, art professor and new media artist Matt Kenyon describes his decades-long pursuit to “make visible” the invisible hand of destructive market forces and exploitative power dynamics, which have been normalized by disinformation-filled narratives of progress. Kenyon builds on the sci-fi trope of the classic horror film *The Invisible Man* (1933), which he offers as an analogue for his own public art program: “Through my artwork, I aim to make the unseen visible, much like the way the fresh snow in *The Invisible Man* reveals the hidden protagonist. I strive to highlight the often unseen influence of power and wealth, bringing these forces into focus by making work that might catch the footprints they leave in the snow.”

In chapter 6, geographer Jessie Poon and epidemiologist Laurene Tumiel Berhalter team up to examine the influence of socioeconomic, demographic, and institutional factors on trust (or lack thereof) in medicine based on data from the General Social Survey (GSS) and specialized literature on the subject. While a survey of twenty-eight countries conducted by the market research company Ipsos (2021) found the medical profession to be “most trustworthy,” the public trust in medical scientists in the United States lags way behind, as only 29 percent of Americans report they have high confidence that members of the medical scientific community will act in the public’s interest (Kennedy, Tyson, and Funk 2022). While vaccine misinformation has no doubt exacerbated the problem in recent years, the historical arch provided by Poon and Tumiel Berhalter suggests that there are deeper long-range issues at play that must be understood. As they write, “The American Association of Medical College’s Center for Health Justice recognizes the importance of trust in the delivery of quality health care for all and the reduction of health inequity by establishing Principles of Trustworthiness as a guide for health care and public health. Trust has been identified in a burgeoning literature as an important mechanism for combating vaccine hesitancy and misinformation (Franic, 2022). Understanding the influences

that explain trust and leveraging existing resources can help inform public policy regarding how trust may be improved to mitigate against misinformation at a time of public health crisis.”

This is precisely the starting point of the final chapter of the collection authored by biochemist Jennifer Anne Surtees. In her essay, Surtees makes the point that public engagement is vital to promote mutual understanding between the medical professionals and the communities they serve: “As researchers at a public university, we have a responsibility to engage with our community . . . provide our citizenry with the tools and knowledge to understand, regulate, ensure ethical and equitable use of, and derive maximum benefit from the astonishing advances in these biomedical sciences.” Surtees herself has focused her own professional activity on hands on, inquiry-based community engagement practices involving K–16 students and adult populations on issues ranging from genome and microbiome literacy to vaccine education. Her stated goal provides a fitting conclusion to the volume: “to nurture robust community partnerships to better understand the unique crosscutting needs of diverse communities and to develop an infrastructure of trust through which to communicate scientific advances, from genomic medicine to pandemic risk.”

As we bring this introduction to a close, Yuval Harari, the historian who authored *Sapiens: A Brief History of Humankind* (2018), Tristan Harris, computer scientist and co-founder of the Center for Humane Technology, and his co-founding partner, mathematician and physicist Aza Raskin, published an opinion piece in the *New York Times*, which includes dire warnings about democracy unless we learn to live with (and protect ourselves from) AI-spread and AI-generated illusions and update our nineteenth-century institutions to cope with twenty-first-century realities:

The specter of being trapped in a world of illusions has haunted humankind much longer than the specter of AI. Soon we will finally come face to face with Descartes’s demon, with Plato’s cave, with the Buddhist Maya. A curtain of illusions could descend over the whole of humanity, and we might never again be able to tear that curtain away—or even realize it is there . . . In social media, primitive AI was used not to create content, but to curate user-generated content. The AI behind our news feeds is still choosing which words, sounds and images reach our retinas and eardrums, based on selecting those that will get the most virality, the most reaction, and the most engagement . . . Democracy is a

conversation, conversation relies on language, and when language itself is hacked the conversation breaks down and democracy becomes untenable. If we wait for the chaos to ensue, it will be too late to remedy it . . . The first step is to buy time to upgrade our 19th-century institutions for a post-A.I. world, and to learn to master A.I. before it masters us. (Harari, Harris, and Raskin, 2023)

We would further argue that the first step in protecting human societies against the curtain of AI-spread and AI-generated illusions must include an educational line of defense so that we can be better prepared for not just what's to come but what is already here. Thus, the urgent upgrade of "our 19th-century institutions" that Harari, Harris, and Raskin call for would need to include a revamping of our educational practices beyond the nineteenth-century version of disciplinary silos. We can no longer afford to be segmented into nearly hermetic fields if we are serious about figuring out how to navigate the AI-assisted present-future. It bears repeating that the urgent challenges outlined in this volume are not mere technical problems that can be effectively solved inside technical fields alone. When and where the vital matter of language and communication is involved, the humanities and arts must be key players, along with their partners in the social sciences, as we try to (re)train ourselves in the "art of reading reality" (Castillo and Egginton 2016, 2021). As our contributing colleagues remind us in their own scholarly practices, the kind of collective (re)training we are advocating here requires a degree of disciplinary self-examination and, most importantly, a willingness to engage with colleagues working in other fields and members of the larger community, outside our own academic circles. After all, communication is a two-way street predicated on trust.

References

- BBC Trending. 2016. "The Saga of 'Pizzagate': The Fake Story That Shows How Conspiracy Theories spread." BBC (Dec 2). www.bbc.com/news/blogs-trending-38156985
- Bisgaard, Martin. 2019. "How Getting the Facts Right Can Fuel Partisan-Motivated Reasoning." *American Journal of Political Science* 63, no. 4: 824–839.
- Castillo, David, and William Egginton. 2016. *Medialogies: Reading Reality in the Age of Inflationary Media*. New York: Bloomsbury Press.

- . 2021. *What Would Cervantes Do? Navigating Post-Truth with Spanish Baroque Literature*. Kingston, ON: McGill-Queen's University Press.
- Cervantes, Miguel de. 2010. *Don Quixote de la Mancha*, 2 vols. Ed. John Jay Allen. Madrid: Cátedra.
- Franic, J. 2022. "What Lies behind Substantial Differences in COVID-19 Vaccination Rates between EU Member States?" *Frontiers in Public Health* 10: 1–15.
- Hanbury, Mary. 2018. "Mattress Firm Responds to the Wild Conspiracy Theory about Its Business That People Are Going Crazy Over." *Business Insider* (Jan 24). www.businessinsider.com/mattress-firm-responds-to-conspiracy-theory-about-its-business-2018-1
- Harari, Yuval. 2018. *Sapiens: A Brief History of Humankind*. New York: Harper.
- Harari, Yuval, et al. 2023. "You Can Have the Blue Pill or the Red Pill, and We're Out of Blue Pills." *New York Times* (March 24). www.nytimes.com/2023/03/24/opinion/yuval-harari-ai-chatgpt.html
- Kennedy, Brian, Alec Tyson, and Cary Funk. 2022. "Americans' Trust in Scientists, Other Groups Declines." Pew Research Center.
- Khodabakhsh, Ali, Raghavendra Ramachandra, Kiran B. Raja, Pankaj Wasnik, and Christopher Bush. 2018. "Fake Face Detection Methods: Can They Be Generalized?" (Conference Paper, September 2018). www.researchgate.net/publication/327350438_Fake_Face_Detection_Methods_Can_They_Be_Generalized
- Lago, F., C. Pasquini, R. Bohme, H. Dumont, V. Goffaux, and G. Boato. 2022. "More Real Than Real: A Study On Human Visual Perception Of Synthetic Faces." *IEEE Signal Processing Magazine* 20, no. 10: 23–40.
- Meta AI. 2020. "Deepfake Detection Challenge Dataset." (Meta, June 25). <https://ai.meta.com/datasets/dfdc>
- Nagle, John Copeland. 2008. "The Evangelical Debate over Climate Change." *U. St. Thomas LJ* 5: 53.
- Pasquale, Frank. 2015. *The Black Box Society: The Secret Algorithms that Control Money and Information*. Cambridge, MA: Harvard University Press.
- Rouvroy, Antoinette, and Thomas Berns. 2013. "Algorithmic Governmentality and Prospects of Emancipation: Disparateness as a Precondition for Individuation through Relationships?" Translated by Elizabeth Libbrecht. *Réseaux* 177, no. 1: 163–196.
- Shane, Tommy, Emily Saltz, and Claire Leibowicz. 2021. "From Deepfakes to TikTok Filters: How Do We Label AI Content?" *First Draft* (May 12).
- Stiegler, Bernard. 2019. *The Age of Disruption: Technology and Madness in Computational Capitalism*. Translated by Daniel Ross. Cambridge: Polity Press.
- Thompson, Stuart A., Tiffany Hsu, and Steven Lee Myers. 2023. "Conservatives Aim to Build a Chatbot of Their Own." *New York Times* (March 22). www.nytimes.com/2023/03/22/business/media/ai-chatbots-right-wing-conservative.html