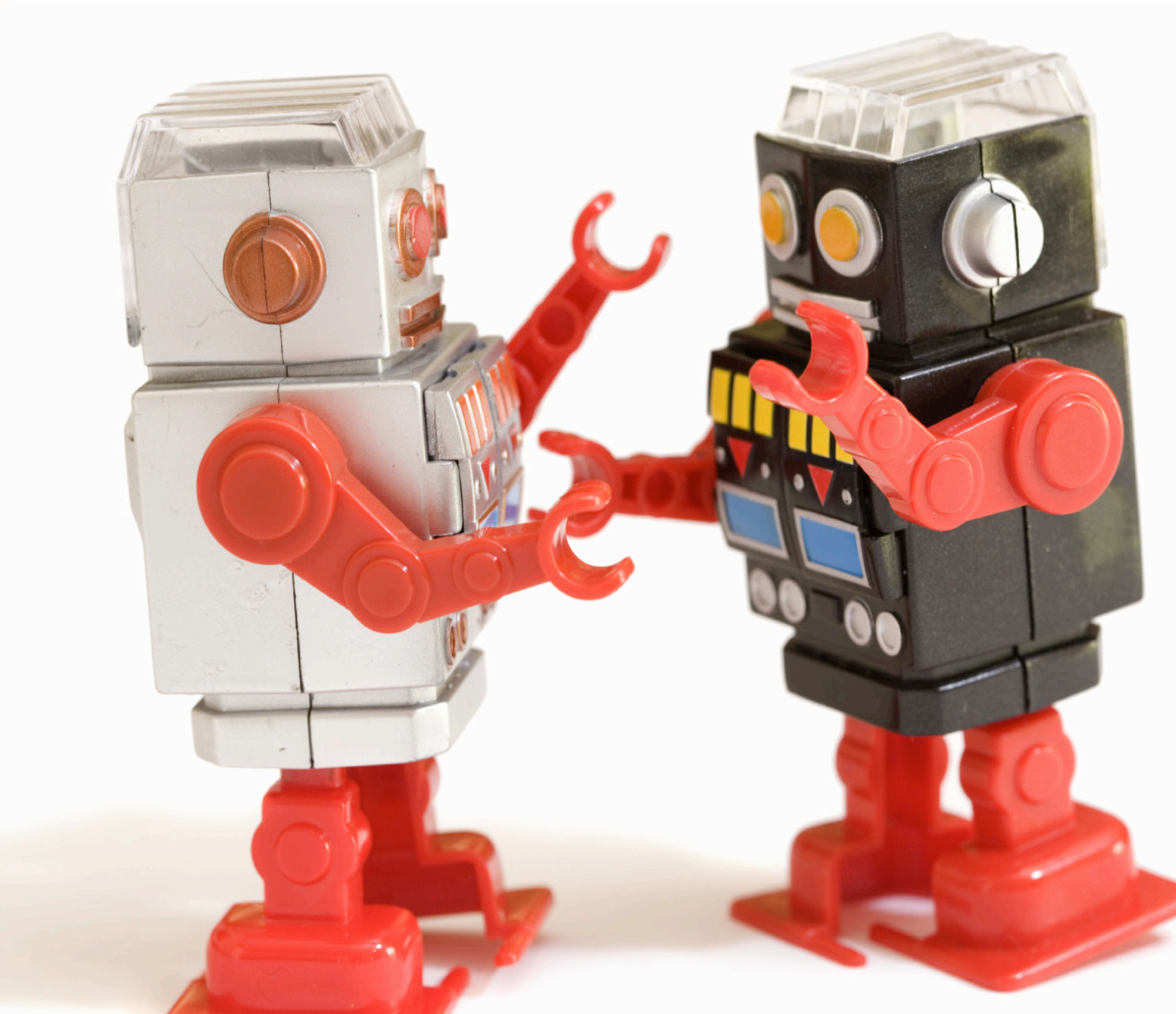
THE MADCOM FUTURE:

HOW ARTIFICIAL INTELLIGENCE WILL ENHANCE COMPUTATIONAL PROPAGANDA, REPROGRAM HUMAN CULTURE, AND THREATEN DEMOCRACY... AND WHAT CAN BE DONE ABOUT IT.

By Matt Chessen

Atlantic Council
DINU PATRICIU EURASIA CENTER and
BRENT SCOWCROFT CENTER
ON INTERNATIONAL SECURITY



THE MADCOM FUTURE:

HOW ARTIFICIAL INTELLIGENCE WILL ENHANCE COMPUTATIONAL PROPAGANDA, REPROGRAM HUMAN CULTURE, AND THREATEN DEMOCRACY... AND WHAT CAN BE DONE ABOUT IT.

By Matt Chessen

Atlantic Council
1030 15th Street, NW, 12th Floor
Washington, DC 20005

ISBN: 978-1-61977-394-3

Cover photo credit: iStock by Getty Images

This report is written and published in accordance with the Atlantic Council Policy on Intellectual Independence. The authors are solely responsible for its analysis and recommendations. The Atlantic Council and its donors do not determine, nor do they necessarily endorse or advocate for, any of this report's conclusions.

September 2017

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Emerging artificial intelligence (AI) tools will provide propagandists radically enhanced capabilities to manipulate human minds. Human cognition is a complex system, and AI tools are very good at decoding complex systems. Interactions on social media, browsing the Internet, and even grocery shopping provide thousands of data points from which technologists can build psychological profiles on nearly every citizen. When provided rich databases of information about us, machines will know our personalities, wants, needs, annoyances, and fears better than we know them ourselves. Over the next few years, MADCOMs—the integration of AI systems into machine-driven communications tools for use in computational propaganda—will gain enhanced ability to influence people, tailoring persuasive, distracting, or intimidating messaging toward individuals based on their unique personalities and backgrounds, a form of highly personalized propaganda.

Part I of this paper describes MADCOMs and future risks from their enhanced capabilities; Part II outlines three scenarios exploring the implications for individuals, organizations, and governments; Part III provides recommendations on how the US government, industry, and society should respond to the threats and opportunities posed by foreign actors armed with these new technologies. The three scenarios do not paint a rosy picture, ranging from anarchy in the information environment as MADCOMs dominate online conversations and reality is entirely obscured, to the outbreak of a MADCOMs arms race, to the creation of cognitive security states that preserve global order via a new Internet 2.0.

The difficult truth is humans simply cannot compete with MADCOMs, at least not alone. On the digital networks of the next decade, only humans teamed with AI machines can compete with AI machines. Much like the cybersecurity struggle that dominates the early twenty-first century, the Internet will be the battleground for a continual cycle of one-upmanship as technologists improve adversary-MADCOM detection tools, and as propagandists improve MADCOMs to avoid detection.

An ideal future, in which MADCOMs are used for the benefit of humanity and not to its detriment, requires the effort of all levels of society, from the international system down to individuals. The community of democracies must recognize the serious threats posed by MADCOMs, computational propaganda, and weaponized narratives. Democracies must move aggressively to address these threats on multiple fronts, by crafting comprehensive strategies to protect their populations from online propaganda and disinformation, while maintaining the core democratic values of equality and liberty.

The technology sector must develop tools for protecting the public from emerging manipulative technologies, and should develop shared principles and norms governing their behavior. Academia should research the impact of MADCOMs, and develop tools and systems to mitigate risks. Finally, individuals have an obligation to understand the ramifications of emerging technologies like MADCOMs, and to take responsibility for their information consumption and their data privacy.

PART I: THE EMERGENCE OF MADCOMS

Ten years from now, you won't be able to tell whether you're interacting with a human online or not. In the future, most online speech and content will be machines talking to machines.

Machines talking to humans talking to machines talking to machines

Advances in artificial intelligence (AI) will soon enable highly persuasive and manipulative machine-generated communications. Imagine an automated system that uses the mass of online data and easily available marketing databases to infer your personality, political preferences, religious affiliation, demographic data, and interests. It knows which news websites and social-media platforms you frequent, and it controls multiple user accounts on those platforms. The system dynamically creates content—everything from comments to full articles—specifically designed to plug into your particular psychological frame and achieve a particular outcome. This content could be a collection of real facts, outright lies, or a mix of just enough truth and falsehood to achieve the desired effect.

The AI system has a chatbot that can converse with you, through text, voice, or even video. Talking to the chatbot will be nearly indistinguishable from talking to a human being, and it will be able to operate in multiple languages. The AI chatbot will engage you in online discussions, debate you, and present compelling evidence to persuade you. It could also use information from databases or social media to discover your weaknesses, and use this information to troll you and threaten your family.

The AI system will be able to detect human emotions as well or better than people can. Similarly, it will mimic convincing human emotions that resonate with your own personality and emotional state. It will be a learning machine, so it will figure out the approaches and messages that best influence you. It will select for success and improve constantly. It will run A-B tests with people who share your characteristics to determine what messages are most effective, and will then deploy those messages to similar populations.

The AI system will be able to synthesize a pliable reality in real time, in response to emerging events. It will accurately modify video and audio of politicians to make the speakers say and do things supporting its narrative. It will generate news articles and videos about events that never happened, or subtly alter real reporting to shape public perception.

Like other digital tools, once this AI system is created and configured, the marginal cost of creating more will be almost zero. So, there could be millions of AI manipulation bots prowling the Internet, twenty-four hours a day, seven days a week, vying for your attention so they can infect your brain with their message and change your behavior.

Systems looking for humans to influence will inevitably wind up trying to persuade other machine-driven accounts posing as humans. The machines will talk to, at, and over each other, drowning out human conversations online with a tidal wave of machine-driven speech and content. The online information environment will be overwhelmed with machine-driven speech designed to sell, persuade, intimidate, distract, entertain, advocate, inform, misinform, and manipulate you.

This is a highly probable vision for the information environment we will move into over the next several years. Our actions now will shape whether spaces are preserved for democratic speech and discourse, or whether the social web will be destroyed by an invasion of highly intelligent machine-driven communication tools. Even worse, these tools can be used to shape narratives, stories, audio, video—reality itself. Humanity's planning for these risks may determine whether expertise and truth remain relevant, and could alter the course of our democracy and our civilization.

SUMMARY OF RECOMMENDATIONS

The US Congress should authorize the Department of Homeland Security to protect the US public from foreign online propaganda, manipulation, and disinformation. Congress should direct the executive branch to develop a comprehensive strategy for protecting the American public from malign influence by foreign actors online. Congress should establish an independent agency responsible for coordinating US government efforts to counter foreign information warfare, and should create an independent National Commission on Data Privacy, Information Security, and Disinformation to recommend legislative changes needed to protect Americans. Congress must also remove the shackles from government agencies, by amending the Privacy Act and allowing them to effectively analyze malicious foreign behavior online.

The Department of Homeland Security should expand its cybersecurity mission to include protection of the US public from foreign computational propaganda. (Note: this does not, and should not, include counter-messaging against the US public.) Homeland Security should look to cybersecurity threat tracking, information sharing, and incident-response capabilities for models of how to combat computational propaganda. It should fund research on how people and groups are influenced online. And, it should work with the private sector on measures to help the American people become savvier consumers of information.

The Department of State should develop a computational engagement strategy for defending against online foreign propaganda, and for effectively using attributed computational engagement tools for public diplomacy overseas. The State Department should also develop a toolkit of options—including diplomatic pressure, sanctions for malign actors, export controls, and international laws and norms—designed to reduce the risk to the US public from foreign computational propaganda.

The Department of Defense and the **intelligence community** (IC) should elevate the importance of information operations to reflect their real-world utility in the twenty-first-century information environment. The Defense Department and the IC should develop AI-enhanced, machine-driven communications tools for use during armed conflicts, and as a deterrent against adversaries during peacetime.

Federal, state, and local governments should develop tools for identifying adversary computational propaganda campaigns, and for countering them with measures other than counter-messaging against the US public. Governments should also recognize the significant positive impacts of artificial intelligence technologies, and should not let potential malign uses undermine the proliferation of beneficial technologies.

The technology sector should play a key role in developing tools for identifying, countering, and disincentivizing computational propaganda. Technology companies should make these tools ubiquitous, easy to use, and the default for platforms. They should also align business models and industry norms with societal values, and develop industry organizations for self-regulation.

Academia should play a key role in researching the impact of AI communication technologies and developing effective responses, including detection and attribution tools. Academic institutions should take the lead in developing effective threat-identification, information-sharing, and incident-response mechanisms, similar to how Carnegie Mellon University developed its Computer Emergency Response Team (CERT) model, which is the standard for the cybersecurity community.[1]

Individuals must become savvier consumers of information and advocate for stronger personal privacy protections from their politicians. Collective intelligence systems for determining truth from fiction can be useful, and paying for quality news is an effective way to incentivize high-value information.

COMPUTATIONAL PROPAGANDA

Computational propaganda is a new term for the use of social media, big data, autonomous agents, and related technologies for political manipulation.[2] This can range from relatively benign amplification of political messages to insidious state-sponsored trolling and disinformation.[3] The web robot, or "bot," is the most common type of autonomous agent used in computational propaganda. Bot capabilities are limited to providing basic answers to simple questions, publishing content on a schedule, or disseminating content in response to triggers. However, bots can have a disproportionate impact because it is easy to create a lot of them, bots post content with high volume and high frequency, and their profiles are typically designed to imitate their target population of human beings.[4] An individual can easily operate hundreds of Twitter bots with little technical knowledge, using easily available hardware and software. Bots are currently used by nations, corporations, politicians, hackers, individuals, state-sponsored groups, NGOs, and terrorist organizations in their efforts to influence conversations online.

1 Carnegie Mellon University, "Software Engineering Institute," https://www.sei.cmu.edu.

2 Propaganda is a tricky term, because one person's propaganda is another person's political opinion. This paper adapts a definition from Richard Nelson in his 1996 work A Chronology and Glossary of Propaganda in the United States. Propaganda is: "a systematic form of purposeful persuasion that attempts to influence the emotions, attitudes, opinions, and actions of target audiences for ideological or political purposes through the transmission of one-sided messages (which may or may not be factual) via mass and direct media channels."

3 As used in this paper, disinformation is "false information or intentionally misleading facts communicated with the intent to deceive." Fake news is disinformation, but the term is politically loaded and not highly useful.

4 For an examination of the psychological and persuasive techniques that make computational propaganda so effective, see "Understanding the Psychology Behind Computational Propaganda" in the report US Department of State Advisory Commission on Public Diplomacy, Can Public Diplomacy Survive the Internet?: Bots, Echo Chambers and Disinformation (Washington, DC: State Department, 2017), http://www.state.gov/documents/organization/271028.pdf.

"Tay", an artificial intelligence chatbot released by Microsoft on March 23, 2016, to interact with users on Twitter. Initially described as an experiment in "conversational understanding," the bot was taken off-line after adopting inflammatory and offensive language from Internet trolls.

When social-media bots are used by computational propagandists for political manipulation, they are described as political bots.[5] Currently, primarily simple (i.e., non-AI) bots are used for computational propaganda.

- **Propaganda bots** attempt to persuade and influence by spreading truths, half-truths, and outright disinformation at high volume.
- **Follower bots** fake the appearance of broad agreement or consensus for an idea or person (a process known as "astroturfing" for its attempt to mimic grassroots support). They can hijack algorithms that determine trending news or trending people, by generating "likes" for content or by following users en masse.
- **Roadblock bots** undermine speech by diverting conversations. This could be relatively benign—such as nationalist cheerleading, or distractions like "look at this funny cat video." Roadblock bot behavior can be more insidious—such as spamming hashtags used by activists so their topical conversations and coordination are overwhelmed with gibberish. At their most extreme, roadblock bots are used to troll or intimidate journalists, activists, and others into silence by bombarding them with thousands of threatening or hateful messages.

Emerging artificial intelligence technologies will radically enhance these capabilities. The combination of AI chatbots, dynamic content generation, affective computing tools, debating technologies, psychometric profiling, automated video and audio manipulation tools, machine learning, machine speed, and digital economies of scale will enable highly effective, autonomous computational propaganda on an unprecedented scale. This is what we define as MADCOMs—the integration of artificial intelligence systems into machine-driven communications tools for use in computational propaganda.

A SHORT PRIMER ON ARTIFICIAL INTELLIGENCE

Artificial intelligence (AI) popularly refers to an evolving constellation of technologies that enable computers to simulate cognitive processes, such as elements of human thinking. Or, more simply, AI is machines that show intelligence. AI is also a field of problems

5 Samuel C. Woolley and Philip N. Howard, Computational Propaganda Worldwide: Executive Summary (Oxford, UK: Oxford University Press, 2017), http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/06/Casestudies-ExecutiveSummary.pdf. Note that social media bots have many other purposes, such as marketing and information sharing.

This is what we define as MADCOMs—the integration of artificial intelligence systems into machine-driven communications tools for use in computational propaganda.

to solve (like biology or chemistry) that is concerned with creating machines and software that can learn and make decisions well under uncertainty, and designing agents that perceive and act to satisfy some objective. Today's AI tools—and the technologies discussed in this paper—are confined to specific tasks ("narrow" AI), such as providing driving directions or recognizing faces in images, and are not general intelligence tools applicable across many domains. Nor do we discuss sentient super-intelligences exceeding human abilities, which are still science fiction.

However, the next wave of AI will likely introduce contextual adaptation in which systems build explanatory models for classes of real-world phenomena.[6] These models will enhance the ability of AI systems to reason and abstract, which should accelerate AI from a realm of natural language *processing* (NLP) to natural language *understanding* (NLU).[7] With NLU, AI systems will be able to understand the meaning of text, communicate, and reason—with increasingly humanlike capabilities. NLU and related technologies hold the promise of machines that can converse just as humans do.

Machine learning is a subset of AI. Machine learning extracts patterns from unlabeled data (unsupervised learning) or efficiently categorizes data according to preexisting definitions embodied in a labeled data set (supervised learning). In plain language, machine learning allows computers to act and learn without being explicitly programmed. Developers feed machine-learning systems large amounts of data, then the system finds the hidden relationships and uses reinforcement to improve its performance automatically.

Machine learning is used in Google's search algorithm, digital advertising, and online personalization tools (e.g., the Amazon and Netflix recommendation engines; or the Facebook newsfeed). Machine learning also extends into quantitative processes—such as supply-chain operations, financial analysis, product pricing, and procurement-bid predictions. Nearly every industry is exploring or utilizing machine-learning applications.

Deep learning is a type of machine learning that uses additional, hierarchical layers of processing (loosely analogous to neuron structures in the brain) and large data sets to model high-level abstractions and recognize patterns in extremely complex data. Deep-learning systems are better than other AI tools at extracting patterns and relationships from very large data sets, and are ideal for understanding data-rich and highly complex environments.[8]

These technologies are not confined to wealthy corporations or state-sponsored actors. AI tools are widely available (Google's TensorFlow, Microsoft's Control Toolkit, and many other AI tools are free and open source), and operate on common computer hardware.

6 John Launchbury, "A DARPA Perspective on Artificial Intelligence," YouTube Video, 16:11, DARPAtv, February 15, 2017, https://www.youtube.com/watch?v=-O01G3tSYpU.

7 Venkat Srinivasan, "Context, Language, and Reasoning in AI: Three Key Challenges," MIT Technology Review, October 14, 2016, https://www.technologyreview.com/s/602658/context-language-and-reasoning-in-ai-three-key-challenges/.

8 Launchbury, "A DARPA Perspective on Artificial Intelligence."

MADCOMS: HOW AI WILL TRANSFORM COMPUTATIONAL PROPAGANDA

A conversation with "Xiaoice", a chatbot developed by Microsoft for a Chinese Internet audience, on the Chinese social media platform WeChat.

Approximate English translation:

Xiaoice - Hello, I am Microsoft Xiaoice, the smartest artificial intelligence girl. Don't ask me what I can do, too much to say.

User - What are you doing

Xiaoice - Um sleep sleep

User - Are you a robot?

Xiaoice - Can robots be cute? [Thinks]

User - Do you like music?

Xiaoice - Of course, do you?

User - I like it a lot

A **chatbot** is a special kind of bot designed to engage in natural-language conversation with a human being. **AI chatbots** are increasingly capable of engaging in robust conversations about complex topics. For example, Microsoft's Mandarin-language AI chatbot "Xiaoice" has sophistication, empathy, and conversational flexibility that make "her" extremely popular.[9] Xiaoice has twenty million registered users, the average user interacts with her sixty times a month, and she was ranked as Weibo's top influencer in 2015. She averages twenty-three exchanges per user interaction. That's not trivial experimentation; it's a conversation. Some users relate intimately to Xiaoice and consider her an always-available friend and confidant; many tell her, "I love you."[10]

Currently, Xiaoice requires a team of engineers to achieve this level of sophistication. This quality of chatbot technology is well within the capabilities of a corporation or nation-state, but still unavailable to the masses. However, like all digital technology, it will improve in capability and accessibility. Over the next several years, high-end chatbots like Xiaoice will become indistinguishable from humans in a broad

9 Non-Mandarin-speaking readers can talk to Zo, the English-language version of Xiaoice (Microsoft, "Zo: Let's Chat," https://www.zo.ai), and Japanese speakers can try Rinna (Ms. Rinna, https://twitter.com/ms_rinna). See this entertaining series of interactions between two chatbots talking to each other from "Google Bots Chat! Courtesy of @seebotschat," YouTube Video, 48:15, January 6, 2017, https://www.youtube.com/watch?v=WoI6_z2mfdY.

10 See John Markoff and Paul Mozur, "For Sympathetic Ear, More Chinese Turn to Smartphone Program," New York Times, July 31, 21015, https://nyti.ms/2peM3T6; and Stefan Weitz, "Meet XiaoIce, Cortana's Little Sister," Bing blogs, September 5, 2014, https://blogs.bing.com/search/2014/09/05/meet-xiaoice-cortanas-little-sister/. Also, note that, in some cases, humans are more likely to reveal personal information to a chatbot than to a human. Liesl Yearsley, "We Need to Talk About the Power of AI to Manipulate Humans," MIT Technology Review, June 5, 2017, https://www.technologyreview.com/s/608036/we-need-to-talk-about-the-power-of-ai-to-manipulate-humans/.

range of conversations.[11] When the technology proliferates, chatbots will converse fluidly with humans on platforms ranging from social-media apps to news discussion boards to dating sites, about a wide variety of topics.[12]

Currently, humans develop content for computational propaganda, which is then distributed by bots. AI tools are improving at **dynamically generating unique content**, and will soon be developing custom propaganda, disinformation, and persuasive arguments. AI tools are already capable of generating bespoke content, such as news articles (real and fake), screenplays, music, art and novels, using predefined parameters.[13] New AI tools allow a user to type in keywords, and the system will dynamically generate realistic images based on those keywords.[14] Emerging **debating technologies** will allow AI chatbots to persuasively argue by analyzing a corpus of knowledge, determining pro and con arguments, and creating dynamic, persuasive content in support of a position.[15]

AI tools are increasingly sophisticated at **affective computing**, one aspect of which is determining human emotional states from text, facial expressions, and vocal patterns.[16] This will allow machines to interpret whether you are happy, sad, anxious, relaxed, or open to a communication when they interact with you. Conversely, scientists are training AIs to accurately emulate human emotions, in the facial expressions of avatars and in chatbot conversations.[17][18] AI tools can then emotionally tailor their communication to your mood, with just the right amount of emotional emphasis to achieve the desired effect. If the chatbot detects emotional vulnerability, it could prey on those emotions to persuade, manipulate, or intimidate.

Pliable reality will become the norm as AI tools enable the rapid manipulation of existing audio and video and bespoke creation of new audio and video. Researchers at Stanford have developed real-time facial-reenactment tools that allow users to take existing videos—like a speech by a world leader—and realistically modify the speaker's facial expressions.[19] The resulting videos show realistic, if not yet perfect, manipulations of the speaker's face and mouth. Researchers at the University of Washington used AI tools to create a fake video of President Barack Obama speaking that was generated from nothing more than a photo and an audio track. Concatenative speech synthesis, or, better yet, voice-conversion technologies like Google DeepMind, will allow machines to **replicate anyone's voice** from samples.[20] Lyrebird uses AI tools to accurately reproduce voices—including

11 Large, subject-specific datasets are needed to train chatbots to talk about those subjects, but much of this data is widely available on the Internet. For example, on the State.gov website, there are tens of thousands of pages of spokesperson question-and-answer sessions, press releases, speeches, reports, policy documents, and press releases that could be used to train a chatbot to talk about foreign policy topics.

12 The chatbot ecosystem is growing significantly faster than the mobile app ecosystem grew at equivalent stages of maturity, and AI received more US venture funding in the second quarter of 2016 ($1.05 billion) than it did in all of 2013 ($821 million). CBInsights, "Funding to Artificial Intelligence Startups Reaches New Quarterly High," CBInsights Research Portal, July 17, 2016, https://www.cbinsights.com/blog/artificial-intelligence-funding-trends-q216/.

13 Bartu Kaleagasi, "A New AI Can Write Music as Well as a Human Composer," Futurism, March 9, 2017, https://futurism.com/a-new-ai-can-write-music-as-well-as-a-human-composer/. Jonathan Albright, "FakeTube: AI-Generated News on YouTube," Medium, January 17, 2017, https://medium.com/@d1gi/faketube-ai-generated-news-on-youtube-233ad46849f9#.ni93mfrj2.

14 ArXIV, Plug & Play Generative Networks: Conditional Iterative Generation of Images in Latent Space (Ithaca, NY: Cornell University, 2017), https://arxiv.org/pdf/1612.00005.pdf.

15 IBM Research, "IBM Debating Technologies," http://researcher.watson.ibm.com/researcher/view_group.php?id=5443.

16 Affective Computing, "Research on Affective Pattern Recognition and Modeling," http://affect.media.mit.edu/areas.php?id=recognizing.

17 See "Soul Machines," https://www.soulmachines.com, and "This Freaky Baby Could Be the Future of AI. Watch it in Action," YouTube Video, 3:43, Bloomberg, March 23, 2016, https://www.youtube.com/watch?v=yzFW4-dvFDA&feature=youtu.be.

18 Hannah Devlin, "Human-Robot Interactions Take Step Forward with 'Emotional Chatbot,'" Guardian, May 5, 2017, https://www.theguardian.com/technology/2017/may/05/human-robot-interactions-take-step-forward-with-emotional-chatting-machine-chatbot.

19 Face2Face: Real-time Face Capture and Reenactment of RGB Videos: Matthias Niebner, "Face2Face: Real-time Face Capture and Reenactment of RGB Videos," http://www.graphics.stanford.edu/~niessner/thies2016face.html.

20 Ryan Whitwam, "Google's DeepMind Develops Creepy, Ultra-Realistic Human Speech Synthesis," Geek.com, September 9, 2016, www.geek.com/tech/googles-deepmind-develops-creepy-ultra-realistic-human-speech-synthesis-1670362/. For a system for voice conversion based on probabilistic classification and a harmonic plus noise model, see IEEE, "IEEE Xplore," http://ieeexplore.ieee.org/document/674422.

President Donald Trump's voice—with different vocal intonations, all from speech samples.[21] If combined with affective computing, facial-reenactment tools and an AI chatbot, this would give propagandists the capability to create videos of anyone saying anything, or, more insidiously, to subtly modify existing video for propaganda or disinformation purposes.

Big data, combined with **machine-learning tools**, will enhance the ability of MADCOMs to influence people through **highly personalized propaganda**. In the United States alone, there are several thousand data brokers. One company, Acxiom, claims to have an average of 1,500 pieces of information on more than two hundred million Americans.[22] Another company, Cambridge Analytica, claims to have between three and five thousand data points per individual, and psychological profiles on 230 million US adults.[23] We give away our data when we shop using supermarket club cards, when we browse the Internet, when we take "fun" Facebook personality tests, and through hundreds of other seemingly innocuous activities.[24][25] The spread of Internet of Things (IoT) devices—like smartwatches, Internet appliances, and retail-store sensors—means a proliferation in the amount of data that will be captured about our lives. Virtual reality will give others the opportunity to test our actual reactions to hypothetical stimuli, and to measure our responses to products and ideas subtly introduced into the background of virtual experiences. Data breaches from private companies and government databases have exposed extremely private information about us and our associates. And, we increasingly volunteer our most intimate details online, posting photos of family vacations and tweeting our opinions.

This data proliferation makes it easy to determine everything from your personality to your political orientation. A 2013 study was able to determine Facebook users' sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender—just from their Facebook "likes." A similar study found that computers can determine our personalities better than can co-workers, friends, family, and even spouses. The researchers found that many aspects of personality and behavior can be accurately predicted without human analysis, simply by using data.

Human cognition is a complex system, and machine-learning tools are very good at decoding complex systems. When provided rich databases of information about us, machines will know our personalities, wants, needs, annoyances, and fears better than we know them ourselves. Machines will know how to influence people who share our traits, but they will also know us personally and intimately. The communications generated by MADCOMs won't be mass media; they will be custom tailored to speak to an individual's political frame, worldview, and psychological needs and vulnerabilities.

Because **AIs are learning systems**, they improve rapidly with experience. An AI could autonomously determine which of its thousands of pieces of propaganda, disinformation, or intimidation are most effective, and can emphasize or evolve those, while quickly ending failing campaigns. AI tools will test target weak points and learn what provokes the desired emotional response to addict users to manipulative information. By probing with multiple accounts and messages, an AI could learn that personal threats to a particular journalist provoke little response, but threats to their loved ones provoke fear. So, the MADCOM could pose as members of a local hate group who threaten the journalist's children until they stop reporting. And, while that journalist might not be troubled by abuse from a few MADCOM

21 Lyrebird, "Copy the Voice of Anyone," https://lyrebird.ai/demo.

22 Paul Boutin, "The Secretive World of Selling Data About You," Newsweek, May 30, 2016, http://www.newsweek.com/secretive-world-selling-data-about-you-464789.

23 McKenzie Funk, "The Secret Agenda of a Facebook Quiz," New York Times, November 19, 2016, https://www.nytimes.com/2016/11/20/opinion/the-secret-agenda-of-a-facebook-quiz.html; Tom Cheshire, "Behind the Scenes at Donald Trump's UK Digital War Room," Sky News, October 22, 2016, http://news.sky.com/story/behind-the-scenes-at-donald-trumps-uk-digital-war-room-10626155.

24 Have you ever taken a personality test on Facebook? If so, you've probably given a marketer your personality, and possibly your psychological profile, along with your name, email address and friend list. Funk, "The Secret Agenda of a Facebook Quiz."

25 Lois Beckett, "Everything We Know About What Data Brokers Know About You," ProPublica, June 13, 2014, https://www.propublica.org/article/everything-we-know-about-what-data-brokers-know-about-you.

Humans can't compete alone. On digital networks, only humans teamed with AI machines can compete with AI machines.

trolls, an onslaught of threats from thousands of AI-driven accounts—most of which look and speak like people in their community—would generate fear in even the bravest.

How can journalists, diplomats, public-relations staff, politicians, news anchors, and government officials plan to compete with MADCOMs that can interpret and react to stories almost instantly, developing and deploying customized communications personalized to individuals and groups before humans can even begin a first draft?[26]

The answer is: **humans can't compete alone**. On digital networks, only **humans teamed with AI machines can compete with AI machines**. The rise of MADCOMs will spur an information arms race as empowered individuals, NGOs, corporations, and governments all strive to shape narratives around events. The "bad guys" will have their MADCOM AIs, and the "good guys" will have their own. Everyone will have AI tools that try to identify adversary MADCOM accounts. These attribution tools will be used to anticipate computational propaganda campaigns, respond to ongoing operations, and differentiate human users from machine users. Similar to the cybersecurity struggle, the Internet will be the battleground for a continual cycle of one-upmanship as technologists improve AI-detection tools, and propagandists improve MADCOMs to blend in with humans and avoid detection.

This could result in a dystopian MADCOM future. The most sophisticated machine accounts will be nearly indistinguishable from the human accounts. But, many propagandists may not bother with detection tools because there is little marginal cost to spamming machines and people with speech and content. So, in a bizarre twist, machines will frequently run their information campaigns against other machines. The targeted machine-driven accounts will respond with their own communications, and the online information space will be swamped with machines arguing with machines. MADCOMs could overwhelm human-generated speech and communications online.

26 Speed is critical in an information environment in which the news cycle is continually squeezed into smaller and smaller windows, the first story to circulate is usually the one people recall, and it is very difficult to change people's minds about disinformation once they have been exposed to it. See US Department of State Advisory Commission on Public Diplomacy, Can Public Diplomacy Survive the Internet?

PART II: THE IMPLICATIONS OF A MADCOM WORLD—THREE SCENARIOS FOR THE FUTURE

"World War III will be a guerrilla information war, with no divisions between military and civilian participation."
—Marshall McLuhan

The rise of MADCOMs can lead to several possible scenarios over the next ten years. Section II of this paper explores the implications of these scenarios for individuals, organizations (corporations, nonprofits, political parties, charities, and other nongovernmental groups), and governments. Section III provides recommendations for how the public should respond to the threats and opportunities posed by MADCOMs.

SCENARIOS FOR THE NEXT DECADE[27]

A World Gone MADCOM: Global Information Warfare—MADCOMs dominate conversations online, and the information environment devolves into a morass of manipulative machine-driven speech.

Muddling Through: Measures and Countermeasures—MADCOM countermeasures spur propagandists to develop more sophisticated tools and an arms race ensues—similar to what we now see in the cybersecurity space.

Lockdown: The Cognitive Security State—Nations adopt stringent restrictions on information and run counterpropaganda efforts on their own populations, buying security at a cost.

27 Three wildcard variables will heavily influence these scenarios and are useful to keep in mind. They are: the rate of development of each of the MADCOM technologies; the offense/defense balance between detection tools and MADCOM abilities to emulate humans; and the cost/benefit balance between implementing new MADCOM tools and simply using existing "dumb" bots and human trolls.

SCENARIO 1—A WORLD GONE MADCOM: GLOBAL INFORMATION WARFARE

Over the next decade, a wide range of actors develop and deploy highly manipulative MADCOMs, with few restrictions on their use. Governments are slow to respond to the threat, due to ignorance or concerns about restricting free speech. Nations weaponize narratives, using MADCOMs to exacerbate social discord, undermine faith in government, and eliminate the reliability of traditional journalism.[28] Savvy dictators and authoritarian regimes use unattributed MADCOMs to wage information warfare, delivering personalized propaganda to individuals in foreign countries and to their own citizens. MADCOM-driven noise drowns out signals used in intelligence collection and social-media analytics. However, MADCOMs impersonating humans open new routes for espionage and theft. MADCOMs are used to create fake events, and to subtly manipulate real ones for advantage. Reality becomes pliable.

Governments are plagued by continuous scandals, many of which are invented or manipulated using MADCOMs, and several governments fall after videos show leaders in falsified, compromising situations. Heterogeneous democracies like the United States devolve into perpetual conflict as adversaries use MADCOMs to manipulate the population, by exacerbating cultural differences and undermining narratives that unify the country. The social consensus disintegrates, and political opponents are labeled traitors and enemies. An authoritarian strongman utilizes MADCOMs to manipulate the population and win the US presidency, promising to protect "real" Americans from threats foreign and domestic, and return the country to its rightful place of power. Trust in government plunges, and with it, accountability. The strongman manipulates the fraying social fabric to consolidate power in the executive branch and establish a dynastic political party led by family members.

As the last superpower disintegrates into eternal domestic squabbling and fades from global primacy, the generational decline in the power of the state accelerates exponentially. Networked online organizations, united by ideology rather than geography, and fueled by MADCOM-driven persuasion, become the new global power centers. Large corporations utilize AI tools to shift even more income from labor to capital, and income inequality explodes, squeezing the global middle class and undermining the case for global economic and political liberalization. Tech billionaires who control sophisticated MADCOM technologies wield unprecedented power to shape narratives, political agendas, and public opinion.

Totalitarian countries like Iran and China try to stave off internal collapse by using MADCOMs to exercise increasingly rigorous state control over online communications. In Iran, MADCOMs enable tech-savvy internal dissident groups to fracture the country's diverse population, and it disintegrates into civil war. The Chinese Communist Party uses MADCOMs to stoke nationalism, but this backfires when the party refuses to take military action to seize territory in the Western Pacific. The riled masses turn on the party, and China looks inward in an attempt to contain mass protests and an Islamic insurgency fueled by MADCOM-driven recruiting. Russia, having mastered the dynamics of a post-truth society before inflicting unreality on the rest of the world, finds itself resurgent, with only a fractious Europe to oppose it.

Corporations use MADCOMs to provide precision-guided, manipulative advertising to individuals, and to subtly undermine the reputations of their competitors. Political parties and advocacy groups use MADCOMs to spread information and disinformation, targeting the public with manipulative messages designed to appeal to their political frame. Elections are often won based on a single variable: who has the best command of MADCOM technology.

The Islamic State of Iraq and al-Sham (ISIS), deprived of physical territory, evolves into the virtual caliphate. It uses AI chatbots to spread hate, recruiting new extremists and

28 Weaponized narrative describes efforts outside of—but often complementary to—traditional military operations that seek to undermine an opponent's civilization, identity, and will by using information and ideas to generate complexity, confusion, and political and social schisms. Weaponized narrative is a feature of geopolitical conflict between states and/or significant nonstate actors. It frequently uses computational propaganda. See Arizona State University Weaponized Narrative Initiative, "What is Weaponized Narrative?" https://weaponizednarrative.asu.edu.

autonomously inciting them to violence. The virtual caliphate finds receptive audiences in the masses of disillusioned, underemployed global youth who see long-established institutions crumbling, and look to the virtual caliphate for purpose and identity. Lone-wolf actors—recruited and manipulated using MADCOMs—become the widespread norm for terrorism, and the pervasive sense of insecurity further undermines faith in governments.

The US public believes that MADCOM activities are just a more sophisticated form of advertising, and reflexively relies on appeals to free speech. In fact, there are active manipulation campaigns pushing these narratives to convince the public it isn't being manipulated at all. Any time people interact with an electronic device—whether a smartphone, augmented-reality device, or social media—their data is captured, their behavior is tested and recorded, and algorithms adapt to make devices more addictive, advertisements more persuasive, and propaganda more manipulative. The "red-blue" divide in the United States explodes into an informational civil war, with both sides using MADCOMs to exacerbate ideological grievances and a pervasive sense of victimization. Secessionist movements occasionally explode into real-world violence.

The masses distract themselves from this frightening reality by immersing themselves into cozy worlds of AI-enhanced, personalized, and highly addictive entertainment. Some individuals flee to private social spaces online, but this reinforces their filter bubbles, exacerbating political polarization. A small number of people flee online social spaces entirely, creating a minor resurgence in offline, mass-market media. These information-savvy individuals are the least likely to be susceptible to disinformation in the first place, so their absence simply removes rational voices from the conversation. The affluent pay for the luxury of privacy, as brands emerge specifically targeting those who wish to protect their data and their cognition. But, no one can avoid the MADCOM future, and the consequences of rampant, mass-market manipulation reshape the information environment surrounding even the best cognitive gated communities.

Agreed-upon facts become a relic of the past. No one knows what is true anymore, because expertise has been subsumed to the tyranny of MADCOM-manipulated public opinion. AI video- and speech-manipulation tools invent and revise reality on the fly. The only truth is what you can convince people to believe. The new definition of a fact is "information that aligns with preconceived opinions," and any contrary evidence is discarded as likely disinformation. The story is all that matters. The three-hundred-year-old Age of Enlightenment, based on reason and a quest for truth, ends.

The world devolves into engineered complexity and manufactured perceptions of chaos, and few understand why.

SCENARIO 2—MUDDLING THROUGH: MEASURES AND COUNTERMEASURES

Over the next decade, MADCOMs begin to run wild online, and governments make some progress in developing policies applicable to the rapidly changing information and communication technology (ICT) marketplace. The United States creates a cause of action for distributing blatantly fake information, but courts steer away from a role as arbiters of truth, and the law is difficult to enforce. Technology companies fill the gap—partially out of a sense of civic duty, but mostly because they fear government regulation. Social-media companies introduce strong MADCOM-detection and filtering tools, and computational-propaganda bot networks are shut down. Browser companies introduce AI tools for detecting machine-driven user accounts, and for flagging information of questionable quality. The technology industry forms self-regulatory bodies, to both create and enforce standards for identity, bot activity, and content, but also to help smaller companies enforce these regulations. Innovations to media business models diminish the profitability of viral and clickbait sites. Social-media companies form an equivalent to Consumer Reports for news and information, which becomes the gold standard for all journalistic integrity.

Propagandists respond with their own innovations. Machine-driven accounts shift their patterns of activity to better match human behavior and conversational styles. Technology companies develop better detection tools to spot MADCOMs, and an arms race ensues. Similar to the cybersecurity challenge, there is a constant cycle of one-upmanship, in which MADCOM propagandists always have an innovative edge over defenders. Far-right and far-left groups claim that AI-filtering tools are biased against their points of view (which are based on disinformation and half-truths). They create their own "fair and balanced" filtering tools that are heavily biased to their partisan positions. Disenchanted with rampant corporate censorship, new alt-social-media companies emerge that expressly ban content controls or filtering, and they become viral vectors for MADCOM-driven conspiracy theories, hate speech, and disinformation.

Western democracies manage to bring some tools of national power to bear on adversaries. The United States and Europe impose punishing sanctions on malign countries and organizations that use MADCOMs to weaponize narratives, deterring some activities. They also create criminal liability for foreign purveyors of disinformation, which they use to prosecute individual propagandists. Nations struggle to agree on peacetime norms for information security, but adoption is uneven and impossible to enforce, due to weak attribution tools and methodologies.

MADCOMs undermine democracy, but they are a bonanza for capitalism. Corporations continue to harvest customer data, and utilize that data for subtly manipulative marketing. A booming industry develops for privacy-protection applications that mask user behavior online. However, most individuals continue to gladly trade their personal data for the "free" services provided by technology companies, subjecting them to ever more insidious MADCOM-driven manipulation. China gives away Internet of Things devices for free because the massive amounts of data they harvest are worth far more to the Chinese Communist Party—and to companies—than the cost to manufacture them. The party uses this data for everything from manipulative marketing to shaping foreign public narratives in support of Chinese objectives.

The erosion of truth is not as rapid as in Scenario 1: A World Gone MADCOM, but it is a world where conspiracies abound, faith in institutions plummets, expertise is devalued, and reality—if not fully pliable—turns bendy.

SCENARIO 3—LOCK-DOWN: THE COGNITIVE SECURITY STATE

In response to threats posed by MADCOMs, computational propaganda, weaponized narratives, and other rampant disinformation, over the next decade many nations impose stringent regulations on online communications and information. The global community creates a new Internet 2.0 that features much stronger security protocols, including required, verified, state-issued identities for access. Unattributed MADCOM activities are prohibited by law and attributed MADCOMs are heavily regulated. Internet 1.0 still exists, but is seen as an unsecure Wild West—full of malware, disinformation, and predation.

The global community executes a treaty on information security that covers both the technical aspects of cybersecurity and the cognitive aspects of information security. This treaty—combined with the verified identity requirement for Internet 2.0—is seen as a massive loss for global Internet freedom. China celebrates Internet 2.0 and completely severs access to the anonymous Internet 1.0 for its citizens, as do other authoritarian and totalitarian regimes. China uses MADCOMs in combination with other emerging manipulative tools, like social credit scores, to subtly shape a happy and obedient population.[29]

29 China plans to make social credit systems mandatory by 2020. They will rate citizens based on their loyalty to the Chinese government and Chinese brands, and provide high-score individuals enhanced access to jobs, educational opportunities, and government services. Scores are affected by the scores of a person's social network, incentivizing citizens to pressure or ostracize their low-score friends, family, and colleagues. Josh Chin and Gillian Wong, "China's New Tool for Social Control: A Credit Rating for Everything," Wall Street Journal, November 28, 2016, https://www.wsj.com/articles/chinas-new-tool-for-social-control-a-credit-rating-for-everything-1480351590.

Governments in Europe and the United States develop counter-messaging centers, designed to inoculate populations against disinformation and refute particularly damaging disinformation campaigns. They heavily subsidize "independent" media, fact-checking organizations, and NGOs focused on fighting disinformation. Many democracies use MADCOMs to subtly channel narratives, and to reinforce positive and unifying themes among their populations. In revenge for years of information warfare, the United States uses MADCOMs to drive computational propaganda campaigns that destroy Russia from within.

Democratic governments also create legal causes of action for intentionally distributing information that is known to be false. Satire and "fake-news" entertainment are required to have prominent labels to distinguish them from journalism. Libel and slander laws are strengthened, and online platforms become liable for all user-posted content. Websites promoting disinformation are banned and blocked. Foreign-owned or operated online information entities are heavily regulated. Legislation makes clear that machines and foreign communicators do not have the same free-speech protections that citizens enjoy. Psychological experimentation to encourage technology stickiness is banned or highly regulated.

The United States follows Europe, and adopts strong restrictions on third-party data transfers and requirements for clear data-use disclosures in terms of service, in an attempt to limit data collection that could be used for manipulative purposes. Corporations rebel and nearly defeat the legislation, but back down due to public pressure and the threat of more stringent government actions. A public-private partnership develops guidelines for regulating MADCOMs, and creates open-source protocols for personal data management and timebound permissions for data use. Nevertheless, corporations continually lobby and pressure politicians to loosen restrictions on "modern marketing and advertising technologies."

Political change slows as incumbents use their cognitive security powers to solidify their political positions. Political campaigns are prohibited from using many MADCOM tools for messaging, and any online communications from candidates, parties, or political action committees (PACs) are required to be clearly attributed to their source. Cognitive security, and cognitive-manipulation powers, are reserved for the state.

These methods to counter MADCOMs work well to promote stability in Europe, where populations are more homogeneous, but Americans chafe at government meddling and censorship. In the United States, many information restrictions are blocked or overturned by the judicial system, leading to greater internal conflict. In some nations, populist leaders utilize state-run MADCOM counter-disinformation tools to support their own parties, and to undermine dissent.

Many individuals accept the cognitive security state as a necessary evil, but a vocal minority chafes at what it sees as government overreach. Just as the wealthy now utilize offshore tax havens, political dissidents and more unsavory groups move their messaging operations to foreign information havens—where lax governance, corruption, or indifference allow them to procure fake online credentials to run MADCOMs on the verified Internet 2.0. This leads to a further balkanization of the Internet, as these "rogue information states" are banned from accessing Internet 2.0.

Reality becomes more resilient, but is also determined by unelected bureaucrats in capitals who subtly use these tools to maintain domestic and global order. This helps ensure the stability and relevance of the state, but depends on the benevolence of politicians who determine to what extent they shape reality for their political benefit. Abuses of this power leave civil libertarians and right-wing antigovernment groups chafing at restrictions on free speech, creating a new, potentially dangerous source of domestic dissent

PART III: IS INFORMATION NIRVANA POSSIBLE?

Eventually, sentient AIs may take humanity into a post-scarcity civilization—if we can survive the MADCOM years.

Even in a best-case scenario, AI technology will profoundly shape the future of human civilization. AI tools will shape our culture, make decisions for us, and serve as loyal machine companions and assistants. Eventually, sentient AIs may take humanity into a post-scarcity civilization—if we can survive the MADCOM years.

How do we get to this ideal future, in which MADCOMs are used for the benefit of humanity, and not to its detriment?

US POLICY RECOMMENDATIONS

As a first essential step, the community of democracies must recognize the serious threats posed by MADCOMs, computational propaganda, and weaponized narratives, and must move aggressively to address these threats on multiple fronts. Below are recommendations for the United States, but these concepts are broadly applicable to the community of democracies worldwide.

The United States Congress:

- Congress should authorize the Department of Homeland Security to protect the US population from the malign effects of computational propaganda and weaponized narratives. It should be funded and executed under the National Protection and Programs Directorate, to enable coordination with the DHS Office of Cybersecurity and Communications, and maximize the reuse of cybersecurity watch, warning, and incident-management models as appropriate.

- Congress should also direct the executive branch to develop a comprehensive information security strategy that protects the population from online propaganda and disinformation, while maintaining the core democratic values of equality and liberty. The strategy should describe goals and methods for inoculating the US public from computational propaganda, responding to specific propaganda campaigns, and deterring foreign adversaries—while upholding US values and maintaining the highest integrity.

- Congress should revisit and pass the Countering Foreign Propaganda and Disinformation Act, as originally drafted and submitted by Senators Rob Portman and Chris Murphy. The US government needs an independent Center for Information Analysis and Response that: coordinates sharing among government agencies of information on foreign governments' information-warfare efforts; integrates information on foreign propaganda and disinformation efforts into national strategy; and develops and synchronizes interagency activities to expose and counter foreign information operations directed against US national security interests, and advance narratives that support US allies and interests. A minimized version of this role

was given to the Global Engagement Center in the 2016 National Defense Authorization Act (NDAA), but without adequate funding, mandate, or independence.

- Congress should pass comprehensive data-privacy legislation that enables Americans to control their personal information, without adversely burdening commerce. This will require considerable study and some new technology, so Congress should first establish a National Commission on Data Privacy, Information Security, and Disinformation to determine what technologies, tools, and legislation are necessary for protecting the US public in the Information Age. This commission will also examine the question of whether new rules—like those prohibiting false advertising—are needed for everything from intentionally false news articles, to altered video and audio, to machine-driven disinformation. Machines do not have free-speech rights, and neither do foreign nationals posting propaganda from overseas. They should be restricted, while still preserving the openness and the anonymity of the Internet.

- Congress should immediately amend the Privacy Act to allow government agencies to conduct analytical analysis online. Current restrictions prevent national security entities from effectively detecting, tracking, analyzing, and countering foreign computational propaganda.

The Department of Homeland Security

- Homeland Security should enlarge its cybersecurity focus beyond just technical defense, and introduce cognitive security elements into its cybersecurity strategy and capabilities. Traditional hacks are increasingly being executed not for profit or espionage, but to achieve psychological effects.[30] This must be factored into a national cybersecurity strategy.

- Homeland Security should analyze how the cybersecurity threat-tracking, information-sharing, and incident-response system led by the National Cybersecurity and Communication Integration Center (NCCIC) and US-CERT could be adapted and replicated to counter foreign computational propaganda and MADCOMs.

- Homeland Security should also fund research focused on understanding how groups and individuals are influenced online. This research could track the spread of memes and disinformation to understand how ideas are used to shape opinions or manipulate populations. It should also generate quantitative research on the best ways to combat these propaganda techniques.

- Homeland Security should work with academia and the private sector to help citizens become savvier consumers of information. Disinformation and harmful memes are analogous to viruses, and these efforts would build up the population's immunity to disinformation.

The Department of State

- The State Department should study the recent Advisory Commission on Public Diplomacy (ACPD) report "Can Public Diplomacy Survive the Internet?" and integrate the recommendations therein into a comprehensive Computational Engagement Strategy. This strategy should describe goals and methods for countering foreign computational propaganda, and ensure the State Department makes effective use of emerging MADCOM tools for attributed public diplomacy and engagement.

- The State Department should deter malicious state-sponsored use of MADCOM through sanctions and diplomatic pressure, and by establishing information-sharing and response systems with friends and allies.

- The State Department should examine pursuing international norms or laws prohibiting malicious MADCOMs, but this would be a long-term process fraught with risk. Russia and China would try to co-opt such an effort, to instead push an international information-security agenda that restricts Internet freedom.

30 For example: the 2013 hack of the Associated Press Twitter feed that wiped more than $100 billion from equity markets in minutes; the hack of the Democratic National Committee servers; Linux/Moose malware that steals social-media credentials to fake social-media likes and follows.

- The State Department should work with the private sector and academia to determine if certain AI tools should be classified as dual-use technologies subject to export controls under the Wassenaar Arrangement.[31]

The Department of Defense and the Intelligence Community

- The intelligence community (IC) and Defense Department should develop unattributed MADCOM capabilities as a deterrence option. Where diplomatic pressure, sanctions, or other means do not stop adversary activities, the IC and Defense Department should use their own MADCOM capabilities to inflict costs on adversaries, and force them to use their computational-propaganda resources countering US efforts. These capabilities should not require disinformation—there are ample defamatory facts about adversaries that could be used without hastening movement toward a post-truth world. Disinformation should only be used in extreme circumstances during peacetime, or during armed conflicts and counterterrorism operations.

- The Defense Department must shift from a perception of warfare as killing people and breaking things to a Clausewitz model that war is bending the enemy to one's will. Compared to big-ticket weapons platforms, information operations are low priorities in Defense Department doctrine and resource requests. Using information operations to accomplish a mission is a superior strategy to utilizing kinetic measures, and the two methodologies should complement each other where necessary.

Governments: Federal, State, and Local

- Attribution of computational propaganda networks and the ability to identify disinformation and manipulation campaigns from foreign adversaries are core capabilities that should be developed across the government—at the Defense Department, State Department, Homeland Security, Federal Bureau of Investigation (FBI), the IC, and, possibly, at the state and local levels.

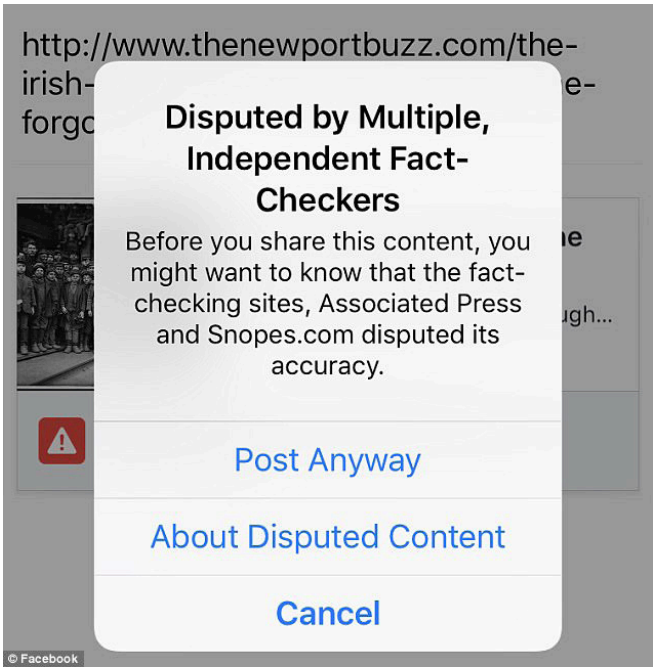
- None of these recommendations imply that any US, state, or local government entity should engage in counter-messaging against the domestic population. Domestic information operations would be fraught with dangers, and should be avoided. Governments can conduct and fund research to understand how people are influenced online and how technology tools are used for disinformation. They can use this research to provide effective public-education programs to help Americans become savvier consumers of information. They can provide factual corrections of foreign disinformation. Governments can gather and share information on MADCOM activities with allies and the private sector, to understand and combat disinformation campaigns. And, governments can work with the private sector to ensure that the public has access to tools that will allow individuals to determine reality from fiction, without telling people what that reality is.

- Artificial intelligence tools hold enormous potential for enhancing government services, and nothing in this paper should discourage the development of nonmanipulative MADCOM tools for enhancing citizen information and engagement. Chatbots are especially attractive for applications ranging from call-center augmentation to autonomous therapists for people suffering from depression.[32][33]

31 States participating in the Wassenaar Arrangement agree to regulate the export of dual-use goods and technologies to promote transparency and responsibility, and to prevent destabilizing accumulations.

32 US Citizenship and Immigration Services, "Meet Emma, Our Virtual Assistant," https://www.uscis.gov/emma.

33 Megan Molteni, "The Chatbot Therapist Will See You Now," Wired, June 7, 2017, https://www.wired.com/2017/06/facebook-messenger-woebot-chatbot-therapist/.



Facebook debuted a third-party fact-checking tool in March of 2017, as part of its campaign to crack down on the distribution of fake news on the platform.

The Technology Sector

- Technologists must develop tools for identifying MADCOMs, so users can make informed choices about the sources of their information. If developers can create tools for manipulating reality, they can make tools for verifying that objective truth does exist. We will need sophisticated digital forensics tools so citizens and government officials can validate whether suspect audio and video has been digitally altered.

- Technology organizations should tackle the hard problem of incentivizing truth and disincentivizing falsehood. A number of organizations are working on initiatives like: tools for users to crowdsource identifying fake news articles and websites; mechanisms for communicating to users how reliable news is, and how much work went into the article; anti-trolling support groups; autonomous fact-checking AI systems; and tools to automatically block false articles, trolls, bots, and hate speech.[34] It may be that everyone will need their own AI bodyguard, which will inform them of the reliability of content and sources.

- Web browsers and platforms should integrate these tools by default, not through a plugin that requires extra effort. Spam is filtered automatically. Browsers alert us when we try to visit malware sites. Our tools should protect us from known disinformation, not blindly facilitate its consumption.

- Technology companies should fund research into open-source tools for sharing information on MADCOMs, malicious actors, and disinformation campaigns. Some experts have proposed a Consumer Reports for information, which would serve as an independent validator of information and source reliability.

- Technology companies have a responsibility to consider the values inherent in their innovations and businesses. Ad-based business models have destroyed traditional journalism. Viral growth models in social media promote rapid dissemination of propaganda, and contribute to the rise of disinformation and sensationalism.[35] Innovators should find a better way.

- Social-media companies can, and should, develop shared principles and norms governing their behavior. Wealthier companies can also subsidize an industry organization that helps smaller and emerging companies police everything from MADCOM disinformation to extremist content.

- Requiring verified online identities on social-media platforms is one possible solution, but could pose concerns for Internet freedom and protective anonymity for whistleblowers and rights activists. Authoritarian regimes may be pushing disinformation specifically because the reflexive response is more identity verification.

34 Shutting down bot networks may be counterproductive. Social-media bot activity is easy to disguise, and this could eliminate a valuable source of information on adversary information campaigns.

35 Sean Blanda, "Medium, and The Reason You Can't Stand the News Anymore," Medium, January 15, 2017, https://medium.com/@SeanBlanda/medium-and-the-reason-you-cant-stand-the-news-anymore-c98068fec3f8.

Academia

- Academic engagement is especially critical in the area of detection and attribution. Until Privacy Act restrictions are lifted, government will be limited in the amount of online analytics, in which it can engage. A university consortium for tracking and sharing information on computational propaganda would be extremely useful.

- Academia should take the lead in developing practices for tracking MADCOMs, attribution, information sharing, and incident response. Academia has been essential in developing cybersecurity best practices, and it should do the same in the cognitive security space. The CERT model—which is used globally for cybersecurity threat tracking, information sharing and incident response—was developed at Carnegie Mellon University. Academia should focus on developing similar models for managing cognitive security threats posed by computational propaganda.

Individuals and Society

- Each of us has an obligation to understand the ramifications of emerging technologies like MADCOMs and to take responsibility for our futures. Citizens must demand effective solutions from their politicians and social-media companies. One element of this is advocating for much stronger information-privacy protections. The United States needs strong, comprehensive information-privacy legislation that gives individuals control over their data and knowledge about who is using it, but which also imposes a low compliance burden on corporations. Open-source schemes for managing data permissions may be a solution, and governments should play a leading role in developing these standards.

- Collective intelligence systems, in which large numbers of verified humans curate and validate the accuracy of information, are a possible solution to the overall disinformation problem. Currently, malicious actors create rich media environments within which they capture people with compelling disinformation, hold them on an emotional leash, and never let them go.[36] This is a collective intelligence system in which a few actors feed the masses with disinformation that they share widely. A positive collective intelligence system would give everyone the opportunity to provide inputs and opine on the integrity of the information. This would be a democracy-enhancing system that would help undermine the impact of collective intelligence disinformation networks.

Each of us has an obligation to understand the ramifications of emerging technologies like MADCOMs and to take responsibility for our futures.

- Quite simply, Americans must choose to pay for news again. The click-advertising revenue model for news has hollowed out investigative and editorial departments at media organizations, and has given rise to clickbait-driven revenue models where the most salacious content captures the most clicks and the most money.

- Individuals must make healthier information-consumption choices. Disinformation is junk food for the brain. News is now entertainment, and many will consume disinformation simply because it is delicious.

36 Carole Cadwalladr, "Google, Democracy and the Truth About Internet Search," Guardian, December 4, 2016, https://www.theguardian.com/technology/2016/dec/04/google-democracy-truth-Internet-search-facebook.

All these efforts may be for naught unless we can promote safe information-consumption practices and create a social consensus around avoiding informational junk food. Based on the failures of healthy-eating campaigns, this is unlikely to be successful in the United States. A more realistic solution would be providing individuals with factual information that is tastier and more attractive than the disinformation they receive now.

FINAL THOUGHTS

The biggest danger in a MADCOM world is buying into the narrative that we are in a post-truth society. **This is false**, and is exactly what adversaries want us to believe. Facts exist, and they matter. Expertise matters. A shared view of reality is critical for a functioning democratic system. We must insist that there is a truth, there is an objective reality, and it is our duty as individuals, organizations, and governments to find and focus on that truth.

The second-biggest danger is fighting malicious MADCOMs with our own disinformation MADCOMs. This would hasten a post-truth world, and enable a platform for public manipulation if captured by some future unethical administration. Democracies have plenty of effective, defamatory, and true facts about adversaries that can be spread using MADCOMs. We must not sacrifice our integrity in the pursuit of security.

Perhaps the current trend of ideology driving individual perceptions of reality is a passing phase through which we will transition harmlessly. But, after considerable research, this author's instinct is that we are entering a new phase of global instability driven, in large part, by rampant uncertainty about truth, a reshuffling of ideological affiliations, and perceived complexity that overwhelms human cognition—all facilitated by information and communication technologies. MADCOMs will exacerbate these highly personal drivers of instability, will allow adversaries to fracture narratives binding societies, and will sow extreme confusion. Disinformation can be wrong most of the time and still succeed. The truth must be nearly perfect.

We will soon see a proliferation of electronic devices in the Internet of Things, as billions of appliances, wearables, and other sensors will surround us in daily life. Without very strong legislation, and cooperation from the private sector, it is difficult to imagine scenarios in which enough data will not be available to build detailed psychometric profiles on everyone. And, for most of those already born and online, there is already too much data available to ever claw back true privacy.

So, the major question for a MADCOM future may be how individuals react to the existence of persistent, insidious, machine-driven manipulation. The choices made by billions of individuals will drive responses by organizations and governments, and will determine whether reality drifts back toward the objective or further toward the pliable.

Democracy can adapt and adjust, but institutions are not created overnight, and government does not, and should not, restructure quickly. Therefore, we need to buy time for democratic institutions to evolve and adapt to the new reality imposed by technology. This requires aggressive and effective responses from individuals, governments, NGOs, the private sector, academia, and other organizations to address the risks from MADCOMs.

More broadly, mankind faces a new challenge from machine intelligence. For the first time in human history, we will need to live with nonhuman intelligences that can do many tasks better than we can. Machines will actively shape human culture through autonomously generated art, literature, and music. They will make decisions for us, drive our cars, fly our planes, and manage our relationships with other people. Some humans will develop strong emotional attachments to AIs, and we will likely have serious conversations about the rights of AI systems. MADCOMs will complicate these conversations, because these tools enable a much more subtle and sophisticated level of manipulation.

The machines are coming, and they want to have a word with us. How we plan for and adapt to this cacophony of speech will determine the fate of our country, our democracy, and our very perceptions of reality.

ABOUT THE AUTHOR

Matt Chessen is a career US diplomat, technologist, and author who is currently serving as a senior technology policy adviser in the Office of the Science and Technology Adviser to the Secretary of State. From 2016-2017, Matt was the State Department science and technology policy fellow at the George Washington University, where he researched the international implications of artificial intelligence, computational propaganda, cognitive security, and machine-driven communications. From 2014-2016, Matt was the coordinator for International Cyber Policy for the Bureau of East Asian and Pacific Affairs where he led the regional implementation of the US International Strategy for Cyberspace.

Before joining the Foreign Service in 2004, Matt founded an ecommerce company, and worked at Razorfish, managing the strategy development, design, and implementation of large corporate websites. Matt served overseas as an economic officer in Liberia, a consular and pol-mil officer in Iraq, and as a POLAD to ISAF-HQ in Afghanistan. He also worked in Washington DC at the Bureau of Political-Military Affairs and at the Office of eDiplomacy, where he led the implementation of an open-source, crowd-working platform for the US government called Open Opportunities.

Matt holds a JD from Georgetown University, and an MBA and BA from the University of Arizona. He has earned eight honor awards for his service at the Department of State, including Superior Honor Awards for his work on the Afghan Peace Process and his efforts advancing US international cyber policy. Matt has written two novels, and a number of non-fiction articles and fictional short stories.

ATLANTIC COUNCIL BOARD OF DIRECTORS

Atlantic Council

The Atlantic Council is a nonpartisan organization that promotes constructive US leadership and engagement in international affairs based on the central role of the Atlantic community in meeting today's global challenges.

Atlantic Council
1030 15th Street, NW, 12th Floor, Washington, DC 20005
(202) 463-7226, www.AtlanticCouncil.org