Disinformation and Entropy

Leveraging AI in the Information Environment

José R. Davis

Considering the Air Force's burgeoning understanding of the information environment in an age of artificial intelligence (AI), effectively leveraging this technology in support of operations in this environment is crucial to success. This article examines the impact of disinformation and potential AI-driven counter-technologies on current and future Air Force operations. Together with improved metrics for assessments of operations, activities, and investments centered on entropy as understood in information theory, a proactive approach to such disinformation and countertechnologies reveals opportunities for the Air Force to win in today's AI era.

Since the term was first coined in the 1970s, *information warfare* has been an amorphous concept, predominantly used by the government and the US military, defined and molded by stakeholders from various backgrounds with different professional vernaculars.[1] Prior to 2017 and the announcement of information as the seventh Joint function, the Department of Defense on the whole had no formal information strategy or information objectives.[2] Similarly, the US Air Force's dispersed information warfare (IW) capabilities had "no comprehensive framework that allow[ed] them to unify their efforts in a way that provide[d] sufficient signal to noise ratio and effective engagement."[3]

Joint doctrine defines the information environment as "the aggregate of social, cultural, linguistic, psychological, technical, and physical factors that affect how humans and automated systems derive meaning from, act upon, and are impacted by information, including the

Captain José Davis, USAF, is a public affairs officer for NATO Allied Air Command, Strategic Communications Division, at Ramstein Air Base, Germany. He holds a master of arts in public administration from Park University.

1. Sandeep Mulgund, Memorandum for C2 of Operations in the Information Environment (OIE) Working Group, Subject: Definitions and Working Descriptions for Information-Related Terms, September 2020; Catherine A. Theohary, "Defense Primer: Operations in the Information Environment," In Focus 10771 (Washington, DC: Congressional Research Service, December 14, 2023), https://crsreports.congress.gov/; and Mike Dahm, "The Reality of War Should Define Information Warfare," *Proceedings* 147, no. 3 (March 2021), https://www.usni.org/.

2. *Joint Warfighting*, Joint Publication (JP) 1-1 (Washington, DC: Chairman of the Joint Chiefs of Staff [CJCS], August 27, 2023), III-25; and Charles R. Grynkewich, "Introducing Information as a Joint Function," *Joint Force Quarterly* 89 (2018), https://ndupress.ndu.edu/.

3. Andrew Caulk, "An Information Warfare Framework for the Department of Defense," *Air & Space Power Journal* 35, no. 1 (April 2021), https://www.airuniversity.af.edu/; and *Public Affairs*, Air Force Doctrine Publication (AFDP) 3-61 (Maxwell AFB, AL: Curtis LeMay Center for Doctrine Development and Education [LeMay Center], September 2020), 12, https://www.doctrine.af.mil/.

individuals, organizations, and systems that collect, process, disseminate, or use information."[4] The lack of a comprehensive framework allowing for unified efforts in this environment made it difficult for the US Air Force to deliver synchronized, practical effects.

Today, the landscape is different. A "complex and volatile global security environment presents profound challenges that erode US global influence and military advantage."[5] Adversaries have become adept at conducting operations below the threshold of armed conflict, which threaten the Department of the Air Force's (DAF) ability to conduct its five core missions—air and space superiority; intelligence, surveillance, and reconnaissance (ISR); rapid global mobility; global strike; and command and control.[6] As a recent RAND report notes, "The role of information and information technologies in strategic competition and military operations has evolved considerably in the first two decades of the 21st century."[7] The challenges of strategic competition are only accelerating with the rapid advancements of artificial intelligence (AI).

On November 17, 2023, the Defense Department released a strategy document on informational power, further codifying terminology and established programs for what is expansively understood as operations in the information environment (OIE), of which information warfare is an adversary-facing component.[8] These operations concern the manner in which information is communicated, transmitted, and processed in the information age. Information—understood as a unified, complex system in which a source pushing a message must overcome noise through a stable conduit to have the desired effect on a receiver—has forced the Air Force's information-related capabilities to become more cross-functional.

Many service functions that contribute to OIE, including public affairs and information operations, have gone through a seismic shift as operators have integrated and collaborated with each other to achieve cohesive effects in the information environment.[9] For example, public affairs, which is responsible for owning public communications and bringing to bear the public personas of institutions into the information environment, has become much more systematic in ensuring its doctrinal mandate of "work[ing] with information operations and strategic communications planners to coordinate and deconflict communication activities."[10]

---

4. *Information in Joint Operations*, JP 3-04 (Washington, DC: CJCS, September 2022), GL-5.

5. Deputy Chief of Staff for Strategy, Integration, and Requirements, US Air Force [USAF], *USAF Operating Concept for Information Warfare*, v1 (Washington, DC: Department of Defense [DoD], March 30, 2022), 2.

6. Michelle Grisé et al., *Rivalry in the Information Sphere: Russian Conceptions of Information Confrontation* (Santa Monica, CA: RAND Corporation, November 2022), 9, https://doi.org/.

7. Grisé et al., 9.

8. "DoD Announces Release of 2023 Strategy for Operations in the Information Environment," DoD, November 17, 2023, https://www.defense.gov/.

9. Mary F. O'Brien, *Integration Imperative: Synchronization of Information Warfare Functions* (Washington, DC: Headquarters US Air Force [HAF], A2/6, February 10, 2021).

10. *Public Affairs*, JP 3-61 (Washington, DC: CJCS, August 2016), I–4.

Today, these various Air Force functions must continue to incorporate and adapt continued advancements in AI, particularly in the realm of generative AI, to provide commanders operational advantage through the information environment.[11] Natural language processing, which attempts to make human communication-like speech and text detectable by computers, and computer vision, which aims to teach computers to act or recommend action on issues based on information gleaned from digital images or other visual input, are and will have a profound impact on OIE, especially in the deployment of large language models (LLMs) and other deep-learning architectures that can masterfully achieve a wide range of tasks, from generating novel text to generating wholly unique images.[12]

Further complicating this situation is the fact that these advances in AI are being employed by Allies and adversaries alike. As a recent NATO report states, the world is entering into a new phase of manipulation in the IE, and "it remains unclear whether, in the long run, defenders or attackers will derive greater benefit from AI systems."[13]

Though general progress has been made in bolstering AI readiness across the service, the Air Force needs to target its AI research and development exclusively on operations for the information environment in order to realize the aim of the service's 2022 *Information Warfare Strategy*—namely, to "deliver automated and AI/ML [machine learning]-enabled tools to support rapid planning and assessment of IW."[14] An analysis of disinformation and AI-based mitigations and the application of entropy as understood by information theory provide options for the Air Force as it looks to win in operations in the information environment.

Operations in the Information Environment and AI

The US Air Force formally defines operations in the information environment as "the sequence of actions that use information to affect behavior by informing audiences; influencing relevant external actors; and affecting information, information networks, and information systems."[15] Further, this understanding of information goes beyond the written or spoken word or even broadcast imagery; it perceives that all activities have a kind of signal that may deliver a message or communicate intent. The IE, for that matter, is more of an "intellectual framework" that assists in comprehending and describing "often-intangible factors" which affect the US military's operational environment.[16]

---

11. JP 3-04.

12. David Morgan, "Using Large Language Models in the DoD Context," DAU [Defense Acquisition University], February 14, 2024, https://media.dau.edu/.

13. Rolf Fredheim, *Virtual Manipulation Brief 2023/1: Generative AI and its Implications for Social Media Analysis* (Riga, Latvia: NATO Strategic Communications Centre of Excellence [StratCom COE], June 2023), 3–12, https://stratcomcoe.org/.

14. *United States Air Force Information Warfare Strategy* (Washington, DC: HAF, July 8, 2022), 6; and Alexander Farrow and Victor Lopez, "AI Readiness in a US Air Force Squadron," *Air & Space Operations Review* 2, no. 2 (2023), https://www.airuniversity.af.edu/.

15. AFDP 3-61, 2.

16. JP 3-04, ix.

In an effort to formalize and integrate Air Force operations in the IE, the Air Force produced the above-referenced *Information Warfare Strategy* and an implementation plan in 2022, merging informational activities and investments across the enterprise.[17] The strategy aims to integrate information across all domains, providing "air component commanders options to modify tempo, timing, and speed of operations."[18]

The service has pushed other initiatives aimed at developing OIE in recent years, in alignment with Joint doctrine.[8] For example, in September 2019 the 16th Air Force became a component numbered air force, making it the only service entity at that level fully focused on information warfare, among its other cyber-related responsibilities.[19] By 2020, the DAF OIE working group had published an official memorandum describing definitions for information-related terms, aimed at clarifying the language used in OIE and providing a consistent lexicon for information-related capabilities.[20]

In 2021, the Air Force merged ISR with its cyber functions, establishing a new directorate postured to synchronize IW-related capabilities.[21] And by 2023, Air Combat Command had become the Air Force's lead major command for organizing, training, and equipping the force for IW.[22] This is only a small sample of recent, myriad initiatives within the Air Force implementing changes and policies for OIE, as outlined by service senior leaders, ranging across doctrine, organization, training, education, leadership, personnel, and policy.[23]

The service's strategy on OIE defines success as the institutionalization and operationalization of informational capabilities across the Air Force. One of the major components to this strategy is providing Airmen advanced tools and systems to deliver IW effects across the competition continuum: "Information Warfare capabilities must be supported by refined analytical methods such as optimization, simulation, decision analysis, artificial intelligence, machine learning, etc."[24]

---

17. *USAF Information Warfare Strategy*; and *USAF Information Warfare Strategy: Implementation Plan* (CUI) (Washington, DC: Headquarters, Department of the US Air Force [DAF], May 2023). Note: the information referenced in the article is not CUI.

18. *USAF Information Warfare Strategy*.

19. Rabia Coombs, "The First Information Warfare Numbered Air Force Welcomes New Commander," Sixteenth Air Force (Air Forces Cyber), July 22, 2022, https://www.16af.af.mil/.

20. "Definitions and Working Descriptions for Information-Related Terms," memorandum, HAF, September 15, 2020.

21. O'Brien, "Integration Imperative," 1–4.

22. "ACC Co-Leads Effort to Hone Information Warfare Readiness," Air Combat Command, March 16, 2022, https://www.acc.af.mil/.

23. Coombs, "First Information Warfare"; O'Brien, "Integration Imperative"; George M. Reynolds, "Achieving Convergence in the Information Environment: Revising the Air Component Structure," *Air & Space Power Journal* 34, no. 4 (2020), https://www.airuniversity.af.edu/; and Sandeep S. Mulgund and Mark D. Kelly, "Command and Control of Operations in the Information Environment: Leading with Information in Operational Planning, Execution, and Assessment," *Air & Space Power Journal* 34, no. 4 (2020), https://www.airuniversity.af.edu/.

24. *USAF Information Warfare Strategy*, 6.

The OIE strategy has two phases: 2022 to 2025, and 2025 to 2029. Each of these phases explicitly directs the leveraging of artificial intelligence for OIE. Phase 1 calls for IW subdiscipline data to be integrated into the Air Force's data fabric, which provides enterprise capabilities that enable the sharing and reuse of data and data tools interconnecting AI users, data, environments, and resources across the Defense Department. Phase 2 calls for the delivery of automated and AI/ML-enabled tools to support rapid planning and assessments of IW.

Importantly, these two phases use whatever tools are available as the technology rapidly advances beyond current capabilities at the time of this writing.[25] One outstanding example of newly developed technology is NIPRGPT—a DoD-approved LLM that can sift through controlled unclassified information documents—developed by Dark Saber, a "software engineering ecosystem" across the Air Force that creates next-generation software capabilities.[26] This and future AI developments will maximize information advantages that ensure the successful employment of airpower in an ever-changing technological landscape.

This article discusses how AI could be leveraged for operations in the information environment, with a special focus on countering disinformation and on OIE assessments, including a new model using information theory.

Inoculating against Disinformation

> *AI will probably have the most impact on information warfare, which could still be highly destructive. We got a glimpse of this when the Russian government interfered with the 2016 presidential election.*
>
> Tom Taulli, *Artificial Intelligence Basics: A Non-Technical Introduction*[27]

The Problem

Disinformation is the deliberate spread of harmful, false, and misleading information.[28] Disinformation is misinformation with a nefarious bent. The most redolent example of disinformation, familiar to many, occurred in 2016 when the Russian government propagated dubious information via social media to manipulate the results of the US presidential election.[29] Disinformation is a plague of the modern information age, exacerbated now by the advancements of generative AI.

25. Morgan, "Large Language Models."
26. "Dark Saber," Dark Saber, accessed May 19, 2024, https://devilops.mil/.
27. Tom Taulli, *Artificial Intelligence Basics: A Non-Technical Introduction* (Berkeley, CA: Apress, 2019), 79.
28. Jon Roozenbeek and Sander van der Linden, *Inoculation Theory and Misinformation* (Riga: NATO StratCom COE, October 2021), https://stratcomcoe.org/.
29. "Grand Jury Indicts 12 Russian Intelligence Officers for Hacking Offenses Related to 2016 Election," US Department of Justice, July 13, 2018, https://www.justice.gov/.

Further, research finds that "people who have been exposed to [disinformation] may continue to rely on it, even if it has been debunked—a phenomenon known as the "continued influence effect."[30] The ability of disinformation to control an individual's cognitive understanding of the world is immensely powerful. And debunking disinformation before it takes root has become exceptionally more difficult with the ability of generative AI to produce manipulative content at scale.

As one RAND report notes, "The world may remember 2022 as the year of generative artificial intelligence: the year that large language models, such as OpenAI's GPT-3, and text-to-image models, such as Stable Diffusion, marked a sea change in the potential for social media manipulation."[31] Moreover, today's AI is tomorrow's least capable AI, as quantum or neuromorphic computing could increase computational power for generative AI. Indeed, US adversaries no longer need to rely on an army of human internet trolls to promulgate disinformation. AI is doing it for them.

In a report by the Center for Countering Digital Hate, researchers discovered AI tools were generating successful images promoting voting disinformation in 59 percent of their tests. These were highly realistic fake images from simple text-based prompts.[32] Further, there is evidence bad actors are using these AI tools now for disinformation. Researchers in the same report saw a drastic upsurge of community notes on X (the platform formally known as Twitter)—for example, user-generated fact-checks added to some posts—by an average of 130 percent per month, demonstrating how disinformation featuring AI-generated images is increasing quickly on social media.[33] In fact, one of the first case studies of voting disinformation, perpetuated by AI and manifesting in campaign videos and automated calls, is playing out at the time of this writing, during the 2024 Indian general election.[34]

Potential Solutions

This concern about disinformation in relation to the rapid advancements of generative AI partially motivated President Joseph R. Biden's Executive Order 14110. The order ensures the safe, secure, and trustworthy development and use of artificial intelligence "by

30. Roozenbeek and van der Linden, *Inoculation Theory*; and Stephan Lewandowsky et al., "Misinformation and Its Correction: Continued Influence and Successful Debiasing," *Psychological Science in the Public Interest* 13, no. 3 (2012): 8, https://doi.org/.

31. William Marcellino et al., *The Rise of Generative AI and the Coming Era of Social Media Manipulation 3.0* (Santa Monica, CA: RAND Corporation, November 2023), 1, https://doi.org/.

32. *Fake Image Factories: How AI Image Generators Threaten Election Integrity and Democracy* (Washington, DC: Center for Countering Digital Hate, March 2024), https://counterhate.com/.

33. *Fake Image Factories*.

34. Meryl Sebastian, "AI and Deepfakes Blur Reality in India Elections," BBC, May 15, 2024, https://www.bbc.com/.

establishing standards and best practices for detecting AI-generated content and authenticating official content." [35] The order directs sweeping actions to protect Americans from the potential risks of AI systems, one of which is deceptive, AI-generated content.

Further, the order directs the US Department of Commerce to develop content authentication and watermarking tools for all federal agencies to use, including the Defense Department. These tools will "make it easy for Americans to know that the communications they receive from their government are authentic."[36] Although this order was recently issued, the US government has been focused on AI ethics and safety dating back to then-President Donald Trump's Executive Order 13859.[37]

Private industry has already begun to experiment with watermarking techniques. In August 2023, Google's DeepMind developed SynthID, which embeds modifications to individual pixels in photos and videos so watermarks are unseen to the human eye, though detectable by computers.[38] Yet in terms of OIE and the continued influence effect, watermarking may be insufficient for curtailing disinformation, largely due to the immense and iterative work needed to make it sufficiently robust, on top of the needed policies to drive its adoption. Even Google has acknowledged that SynthID is "not foolproof against extreme image manipulation."[39]

As Massachusetts Institute of Technology (MIT) Professor Aleksander Madry stated in his testimony before Congress, "We need to start to be more wary than ever about how information reaches us, its trustworthiness and its ability to persuade us."[40] This call for vigilance is heightened in the context of OIE, as DAF equities focus on combating the spread of disinformation.

Artificial intelligence researchers at MIT have developed various techniques that make an image resistant to AI-powered manipulation by adding to the image a carefully crafted, imperceptible perturbation—a small modification in pixels picked up only by a computer.[41] Inoculating an image not only prevents an AI model from trying to manipulate it, but also stymies the spread of disinformation by prebunking it. This approach is twofold, technical in disrupting generative models and psychological in preempting disinformation before it can spread.

---

35. "Fact Sheet: President Biden Issues Executive Order on Safe, Secure, and Trustworthy Artificial Intelligence," White House, October 30, 2023, https://www.whitehouse.gov/; and Exec. Ord. No. 14110, 88 Fed. Reg. 75191 (October 23, 2023).

36. Exec. Ord. No. 14110.

37. Exec. Ord. No. 13859, 84 Fed. Reg. 3967 (February 11, 2019), https://trumpwhitehouse.archives.gov/.

38. Tom Gerken and Philippa Wain, "Google Tests Watermark to Identify AI Images," BBC, August 29, 2023, https://www.bbc.com/.

39. Gerken and Wain.

40. *Artificial Intelligence and Human Rights, Hearing before the Subcommittee on Human Rights and the Law* (testimony of Aleksander Madry, Cadence Design Systems professor of computing, Massachusetts Institute of Technology [MIT]), 118th Cong. (2023), https://www.judiciary.senate.gov/.

41. Rachel Gordon, "Using AI to Protect against AI Image Manipulation," MIT News, July 31, 2023, https://news.mit.edu/.

In the psychological sense, the idea of prebunking and creating a "vaccine" against disinformation derives from a 1960s framework called inoculation theory, advanced by social psychologist William McGuire.[42] Inoculation theory holds that "by exposing individuals to a persuasive message that contains weakened arguments against an established attitude (e.g., a two-sided message, or a message that presents both counterarguments and refutations of those counterarguments), individuals would develop resistance against stronger, future persuasive attacks."[43] Researchers applied this inoculation theory in 2017 within the context of online misinformation.[44] Studies have shown that both partial and full inoculation are effective at countering the effects of misinformation exposure.[45]

The virtue of this approach is in its forced exercise of individuals' rational faculties, allowing them to resist disinformation freely and in their own time, very much like a body's immune system resists a virus on its own after a benign exposure from an immunization. Rational deliberation and the encouragement of people to think through information foster accurate belief formation, allaying the development of partisan bias and susceptibility to misinformation.[46]

Prebunking of and immunization from misinformation have also been seen in current events, when the Biden administration publicly released intelligence information of Russia's various military activities and mobilization throughout the fall and winter of 2021, warning of Russia's building aggression leading up to its February 2022 invasion of Ukraine.

The technical potential of nefarious generative AI could neutralize the potency of prebunking. Fortunately, generative AI can also be employed defensively, in the same way as inoculation theory is used socially. One MIT study funded by the DAF-MIT AI Accelerator program proposes using the MIT-developed AI technique referenced above, dubbed PhotoGuard, which immunizes images and video against the power of diffusion models' ability to manipulate content.[47]

Diffusion models have emerged as impressive tools for generating realistic images, currently surpassing the quality of other image-generating models such as generative adversarial networks. Using a stochastic differential process—which uses random vari-

42. William J. McGuire, "A Vaccine for Brainwash," *Psychology Today* 3, no. 9 (1970).

43. Josh Compton, Ben Jackson, and James A. Dimmock, "Persuading Others to Avoid Persuasion: Inoculation Theory and Resistant Health Attitudes," *Frontiers in Psychology* 7 (2016): 2, https://doi.org/.

44. Sander van der Linden et al., "Inoculating the Public against Misinformation about Climate Change," *Global Challenges* 1, no. 2 (February 2017), https://doi.org/; and Stephan Lewandowsky and Sander van der Linden, "Countering Misinformation and Fake News through Inoculation and Prebunking," *European Review of Social Psychology* 32, no. 2 (2021), https://doi.org/.

45. van der Linden et al.; and Meghan Fitzpatrick, Ritu Gill, and Jennifer F. Giles, "Information Warfare: Lessons in Inoculation to Disinformation," *Parameters* 52, no. 1 (2022): 111, https://press.armywarcollege.edu/.

46. Bence Bago, David G. Rand, and Gordon Pennycook, "Fake News, Fast and Slow: Deliberation Reduces Belief in False (But Not True) News Headlines," *Journal of Experimental Psychology* 149, no. 8 (August 2020), https://doi.org/.

47. Hadi Salman et al., "Raising the Cost of Malicious AI-powered Image Editing," in *PMLR: Proceedings of Machine Language Research* 202 (2023), https://doi.org/.

ables—diffusion models excel in generating and editing images using textual prompts, such as that offered by DALL-E, Stable Diffusion, and Midjourney.[48]

The study mentioned above focused on latent diffusion models, which differ from standard diffusion models mainly in encoding the input image. This approach leverages adversarial perturbations to immunize images, forcing the latent diffusion models to generate images unrelated to the original immunized-input images, demonstrating the ability to immunize images from becoming deepfakes. The study's quantitative results employing PhotoGuard showed success in generating noticeably different images between immunized images and nonimmunized images. Just as inoculation theory in the social sense provides a degree of protection from disinformation, the researchers thus demonstrated AI can provide a degree of protection to content itself from being used for disinformation and deepfakes: "In this paradigm, people can thus continue to share their (immunized) images as usual, while getting a layer of protection against undesirable manipulation."[49]

CariNet is another model developed to provide inoculation against disinformation.[50] Also developed at MIT, CariNet is a novel, semi-supervised artifact attention module that amplifies artifacts—distortions or unwanted features introduced into an image or video during processing—in deepfake imagery to make them more detectable by people. Artifacts in deepfakes vary depending on the technology and methods used to create them. For example, an artifact can be in a manipulated video due to inconsistencies in frame rates, or the speed at which an image is shown, in which a deepfake may not perfectly match the frame rate of the original video, causing stuttering or unnatural movements. CariNet generates "deepfake caricatures"—that is, distorted versions of deepfakes—which magnify unnatural movements in imagery caused by artifacts, hence making them obviously apparent to the human eye.

Importantly, the researchers in various experiments found that exposing deepfakes by amplifying artifacts increases detection rates by people, more so than text-labeled warnings of a deepfake. Moreover, CariNet empowers individuals to exercise their own judgment on the trustworthiness of an image, as opposed to a forced denouncement from a label. Empowering individual judgment strengthens one's immune system against the virus of disinformation: "A system which allows humans to directly detect if a video is doctored will empower them to assess for themselves whether to trust the video."[51]

The engineers of both CariNet and PhotoGuard emphasize the necessity of continued cooperation between developers of these preventative deepfake models and those entities that are determined to curtail the spread of disinformation, such as the US government.[52]

---

48. Gundars Bergmanis-Korāts et al., *AI in Support of StratCom Capabilities* (Riga: NATO StratCom COE, January 2024), 43, https://stratcomcoe.org/.

49. Salman et al., "Raising the Cost," 1.

50. Camilo Fosco et al., "Deepfake Caricatures: Amplifying Attention to Artifacts Increases Deepfake Detection by Humans and Machines," arXiv, Cornell University, last revised April 10, 2023, https://doi.org/.

51. Fosco et al., 9.

52. Salman et al., "Raising the Cost," 8.

Adversaries and nefarious agents could invest in building their own models or upgrading current models, which could make PhotoGuard or CariNet obsolete. Employing these preventative models is not a one-off action, but a matter of continuous development as an element of IW above and below the threshold of armed conflict, whether with PhotoGuard, CariNet, or similar models and research.[53]

Operational advantage or disadvantage is clearly driven by information. The 2004 Abu Ghraib scandal showed how the power of photographs and information could severely hamper US military operations, as insurgents used the imagery as a propaganda tool to fuel greater Iraqi rage and resistance.[54] In a future 2034, AI-generated scandalous and utterly fabricated imagery of US forces could potentially be widely circulated and disseminated, say of an F-35's indiscriminate targeting of civilians, posing a possible serious threat to US military operations. AI's ability to unravel the dichotomy of fact or fiction will undermine airpower.

Fortunately, as discussed, generative AI can be used on the right side as well. The Defense Visual Information Distribution Service (DVIDS), which makes available real-time, broadcast-quality video and still images to media sources, offers one example of application.[55] But uploaded content is also accessible to those who intend to sow disinformation. Yet whether by adding perturbations or amplifying artifacts, DVIDS content could be immunized against mal-intended generative AI, thereby preventing disinformation from taking root.

Entropy, Information, and Assessments

As the Air Force coalesces around a common understanding of OIE, across information-related capabilities and information warfare capabilities, a problem remains. How does the Air Force measure the effectiveness of its operations, activities, and investments in the information environment, across the spectrum of its contributions to the Joint force? To add to the challenge, much of this activity in today's age of strategic competition occurs below the level of armed conflict. Tying action to outcome in the information space is complex and not well-understood, making assessments more challenging to execute successfully.[56]

Current research is developing novel approaches to address this challenge of tying action to information in assessing the IE. One approach advocates perception analysis. Leveraging a review of current literature and interviews with US Air Forces Europe–Air Forces Africa and Pacific Air Forces subject matter experts, one analysis proposes a perception assessment

53. Melissa Heikkilä, "This New Data Poisoning Tool Lets Artists Fight Back against Generative AI," *MIT Technology Review*, October 23, 2023, https://www.technologyreview.com/.

54. Lene Hansen, "How Images Make World Politics: International Icons and the Case of Abu Ghraib," *Review of International Studies* 41, no. 2 (2015), https://doi.org/.

55. "About DVIDS," DVIDS [Defense Visual Information Distribution Service], accessed May 22, 2024, https://www.dvidshub.net/.

56. Katherine A. Batterton, "Operation Assessment in Strategic Competition: Measuring Chinese Communist Party Perceptions" (seminar thesis, Air University, Maxwell AFB, AL, April 2023), 16.

framework "encompass[ed] in three integral components: attention intensity, image/sentiment, and thematic/issue dimensions."[57] This framework takes as a given that a specific message sent may not be interpreted as intended by the receiver. The potential of this approach is in its fundamental acknowledgment of the "complex, nonlinear, interactive, and unpredictable nature of social human interactions."[58]

Another approach to assessments of the OIE called influence quantification (IQ) is purpose built for detection of disinformation narratives. The IQ framework employs "scalable, accurate, and automated discriminants to identify covert foreign influence early in the IE."[59] These discriminants are unusual behaviors or trends that help weed out nefarious actors. This approach is built on network causal inference—that is, in measuring the influence of a source spewing information. Influence quantification can quantify the spread of sentiment through narrative formulation and detection, providing information-related capabilities and IW practitioners meaningful measures of effectiveness.[60]

At the core of IQ's narrative detection is its employment of transformer-based natural language processing for semantic clustering in the IE. By using AI to cluster and sift through copious amounts of data in the form of natural language, IQ can then calculate a causal influence score identifying key influencers propagating information in a network. Interestingly enough, not only can IQ be used defensively for combating disinformation, but it can also be employed offensively for strategic communications.[61]

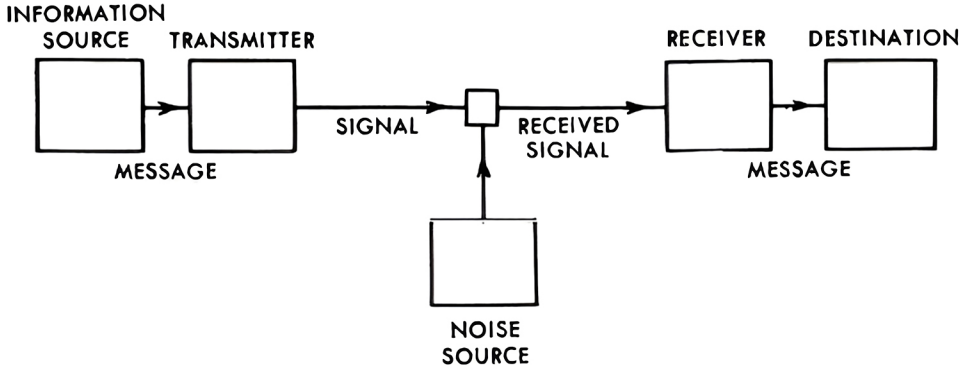

Figure 1. A mathematical model of communication

57. Batterton, 16.

58. Batterton, 8.

59. "Counter Influence Operations Using AI and Causal Inference, with Ethical Considerations," Recent Advances in AI for National Security (RAAINS) Conference, MIT Lincoln Laboratory, Lexington, MA, November 15–17, 2022.

60. Edward Kao, "Influence Quantification for Counter-IO" (slide presentation, RAAINS Conference, November 2022).

61. Stephen T. Smith et al., "Automatic Detection of Influential Actors in Disinformation Networks," *Proceedings of the National Academy of Sciences* 118, no. 4 (2021), https://doi.org/.

Further novel approaches are needed to develop assessment metrics for OIE that are common and useful across information-related and IW capabilities. Surely, each functional capability and every information-related capability have their own measures of success within their respective silos. But as silos are torn down to enhance OIE, IW capabilities need a common metric to measure their efforts toward singular operations, activities, and investments, such as any large-scale exercises the Air Force conducts.[62] In that vein, this article proposes a new metric for the IE, in the form of entropy, as defined under information theory.

Such entropy is the measurement of information as uncertainty or randomness in a channel or a system. The idea of conjoining entropy and information was pioneered by Claude Shannon in *The Mathematical Theory of Communication*.[63] Working at Bell Labs to enhance telecommunications, Shannon was keenly interested in engineering a communications system that could effectively transmit a message through a medium, despite noise and other obfuscations. As a communications engineer, he focused on the technical problem of communication, free of a message's semantic import or connotations, and wanted to mathematically measure how a message could be successfully transmitted. This was the impetus for the development of his idea of entropy as uncertainty.

IE complexity is directly related to the variety of activities present within a complex system. This complexity may be represented numerically using Shannon's entropy formula:

$$H(P) = -XP(i)log2P(i)$$

In sum, the variables calculate the expected amount of information (or uncertainty) in a probability distribution, considering all possible events and their respective probabilities. Stated in another way, "with equally likely events there is more choice, or uncertainty, when there are more possible events."[64]

More information yields greater uncertainty, more choice, and hence more entropy. With increased entropy, there is a greater need for redundancy in a channel, especially in terms of effectively communicating a message. Ultimately, with increase of information entropy, complexity scales to increase randomness to the point of collapse of the signal—or in terms of OIE, the inability to communicate a message effectively.

Shannon used natural language as an example of entropy in information theory. The redundancy of ordinary English is approximately 50 percent, which means half of English is determined by the structure of the language—for example, grammar—and half is chosen freely. Think crosswords or Wheel of Fortune or even a mobile phone's autocomplete text function: The more letters obtained, fewer choices become available, as one homes in on a handful of words.

62. JP 3-04, 115.

63. Claude E. Shannon and Warren Weaver, *The Mathematical Theory of Communication*, Illini Books ed. (Urbana: University of Illinois Press, 1963).

64. Shannon and Weaver, 49.

Entropy as a metric has been shown to be an effective and informative tool for various types of researchers across myriad fields.[65] For instance, cross entropy is used as a loss function in deep neural networks to adjust model weights during training, increasing the accuracy of the model's outputs. Entropy has also been used as a measure of behavioral regularity in studies "uncover[ing] the intricate relationship between habit formations and digital routines," specifically social media habits.[66]

One such study looks at how habits manifest in the digital space, validating the entropy metric as effective in predicting long-term behavior.[67] Within a defense context, entropy has been applied by special operations forces to improve weighting schemes for ranking terrorists during target analysis.[68] In the same sense, entropy can be applied as a common metric for OIE, providing practitioners a quantifiable and predictive way to measure the IE.

Apart from communicating a message, as depicted in figure 1, entropy applies to OIE because the information environment is a complex system. Due to higher entropy, highly complex systems require more cognitive effort to manage and are more cognitively stressful for system participants who are pursuing goals.[69] In this way, cognitive imbalances or disparities by system participants may be inferred from system entropy measures.

For example, information warfare capabilities such as public affairs or information operations could use the information-theory-based notion of entropy to inform their communications strategy by way of observing the IE: If the system in a current state has less information—that is, less entropy—then hypothetically it is the most optimal time for communicating to key audiences. Conversely, for IW capabilities such as weather or ISR, systems with more entropy are optimal because of the wealth of intel and information that can be reaped. This is an example of how entropy as a common metric for OIE can be cross-functional across IW capabilities and information-related capabilities.

Similar benefits could be gained from this metric in terms of key leader engagements. These engagements help commanders create effects in the information environment that can result in a decisive advantage over adversaries and gain rewarding opportunities with Allies and partners. The more complex systems become—that is, the more entropy present—the more likely it is that humans deploy simplifying heuristics. For example, when there is too much information circulating in the information environment, it is hypothetically more effective for the senior leader to keep their messages and engagements simple and short, so they gain better traction toward achieving key-leader-engagement

65. Simon DeDeo et al., "Bootstrap Methods for the Empirical Study of Decision-Making and Information Flows in Social Systems," *Entropy* 15, no. 6 (2013), https://doi.org/.

66. Amir Tohidi Kalorazi, "Habit Formation and Political Persuasion: A Behavioral and Statistical Approach" (PhD dissertation, MIT, September 2023), 3, https://lids.mit.edu/.

67. Kalorazi.

68. William P. Fox et al., "Using the Entropy Weighting Scheme in Military Decision Making," *Journal of Defense Modeling and Simulation* 17, no. 4 (2020), https://doi.org/.

69. Brian Russell and John Bicknell, *The Coin of the Realm: Understanding and Predicting Relative System Behavior*, white paper, Information Professionals Organization, January 24, 2023, https://information-professionals.org/.

goals. Entropy, in this case, could help improve key-leader-engagement timing and improve engagement dossiers to maximize opportunity for favorable outcomes.

Research shows that human beings can only manage so much information. Too much information—too much entropy—can lead to confusion or disorder.[70] Applying Occam's razor—the principle that the simple explanation is preferred to the more complex—to states of higher entropy could be beneficial for public affairs or information operations, where reducing complexity could improve messaging.

Measuring entropy in the information environment requires a lot of data. Fortunately, advancements in machine learning and scalable data-processing systems have made this possible. Using the wealth of data from the Global Database of Events, Language, and Tone (GDELT) project could be a solution. Supported by Google Jigsaw, GDELT "monitors the world's broadcast, print, and web news from nearly every corner of every country in over 100 languages and identifies the people, locations, organizations, themes, sources, emotions, counts, quotes, images and events driving our global society every second of every day, creating a free open platform for computing on the entire world."[71] Using GDELT, entropy could be measured within the IE, which could be useful for information-related capabilities in gauging their impact supporting Air Force operations, activities, and investments.

As nascent as operations in the information environment are, a challenge exists in effectively assessing the expanse of the IE, across the spectrum of information-related and information warfare capabilities. Artificial intelligence can help the Air Force overcome this challenge. Assessing the IE is imperative and a priority of the Joint force.[72] Entropy as a metric—coupled with machine-learning models that can rapidly assimilate the surfeit of open-source information in the mediascape—is one example of a metric that IW and information-related capabilities could use to assess their impact before, during, and after military operations, activities, and investments.

Conclusion

Information acts upon the sociopolitical structures of nation-states in profound ways. With this in mind, the Defense Department takes operations in the information environment seriously and strategically, as information directly impacts commanders' operational environments and the employment of kinetic forces. With rapid advancement, AI too will affect society in profound ways. As one AI expert contends, the potential and problem of

70. John Bicknell and Martin Jetton, *Cognitive Arbitrage: Complexity, Variety and Human Cognitive States Are Related*, white paper, Information Professionals Organization, December 6, 2023, https://information-professionals.org/.

71. "Watching Our World Unfold," GDELT [Global Database of Events, Language, and Tone] Project, accessed July 17, 2024, https://www.gdeltproject.org/.

72. JP 3-04, VI-1.

artificial intelligence is not only one of technology, but also of society.[73] Today's concern is about AI's ability to generate deepfakes and promulgate disinformation. But tomorrow's concern may be related to AI's ability to create real relationships with human beings, whatever that may entail.

As some computer science researchers and humanists have argued, "computer systems designed explicitly to exhibit human-like intentionality (seeming to be about and directed toward the world) represent a phenomenon of increasing cultural importance."[74] AI is often seen through a technical lens, with all the underlying algorithms and engineering of data involved. This view is especially prominent in the Air Force, considering the service's bent toward technocracy. But as some researchers propose, it is time to adopt a humanistic framework of AI that seriously considers how society should interpret a machine's emerging ability to signal intentionality in its actions and behaviors, beyond just chalking up mistakes to generative AI's propensity to hallucinate.[75]

As the Department of the Air Force directs AI research specifically on OIE, it should adopt a framework that converges technical prowess and societal impact. The AI tools for propagating disinformation are becoming dangerously more sophisticated, while the means of combating AI disinformation is increasingly a critical social responsibility not just a technical problem.[76] AI has the potential to propel or pulverize informational advantage for the Joint force. Evidence is clear that disinformation harms the US military's ability to leverage the IE for operational advantage. Furthermore, information's impact on operations needs to be measured: implementing metrics based on entropy as understood by information theory could be one of those measures. In this way, the military—and specifically, the US Air Force—can more effectively collaborate across functions and capabilities as it conducts information warfare in an age of AI. ✈

73. Alger Fraley, *The Artificial Intelligence and Generative AI Bible: From Understanding the Basics to Delving into GANs, NLP, Prompts, Deep Learning, and Ethics of AI* (New York: AlgoRay Publishing, 2023).

74. Jichen Zhu and D. Fox Harrell, "Narrating System Intentionality: Copycat and the Artificial Intelligence Hermeneutic Network," *Leonardo Electronic Almanac* 17, no. 2 (2012): 160, https://groups.csail.mit.edu/.

75. Zhu and Harrell.

76. Henry Kissinger, Eric Schmidt, and Daniel Huttenlocher, *The Age of AI and Our Human Future* (New York: Little, Brown, and Company, 2021), Kindle ed., 96.