The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism

Major Christopher Telley, USA
The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism

by

Major Christopher Telley, USA
AN INSTITUTE OF LAND WARFARE PAPER

The purpose of the Institute of Land Warfare is to extend the educational work of AUSA by sponsoring scholarly publications, to include books, monographs and essays on key defense issues, as well as workshops and symposia. A work selected for publication as a Land Warfare Paper represents research by the author which, in the opinion of the Institute’s editorial board, will contribute to a better understanding of a particular defense or national security issue. Publication as an Institute of Land Warfare Paper does not indicate that the Association of the United States Army agrees with everything in the paper but does suggest that the Association believes the paper will stimulate the thinking of AUSA members and others concerned about important defense issues.

LAND WARFARE PAPER No. 121, October 2018

The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism

by Major Christopher Telley, USA

Major Christopher Telley, an Army Information Operations officer, is assigned to the Naval Postgraduate School (NPS) where he is studying defense innovation and Information Strategy and Political Warfare. He enlisted as a Stinger gunner, serving in Iraq 2004–2005. After commissioning, Telley spent two years as an insurgent platoon leader at the National Training Center. He commanded a reconnaissance troop in Afghanistan 2012–2013, then worked on advanced technology integration at Nellis Air Force Base and around the Pacific. After completing the Information Operations Qualification Course, Telley managed strategic communication for U.S. Army Japan; he led the unit’s cyber campaign, conducted trilateral negotiations between Japan and South Korea and coordinated Army level talks. He holds a Bachelor’s of Science in Political Science from the University of North Georgia.

While assigned to NPS, Telley has worked on projects for Special Operations Command Pacific, Special Operations Command Africa and U.S. Army Pacific. His thesis research—on resistance movements as small state deterrence—was sponsored by the Norwegian Defence Research Establishment.

This paper represents the opinions of the author and should not be taken to represent the views of the Department of the Army, the Department of Defense, the United States government, the Institute of Land Warfare, the Association of the United States Army or its members.

© Copyright 2018 by
The Association of the United States Army
All rights reserved.

Inquiries regarding this and future Land Warfare Papers should be directed to: Director, AUSA’s Institute of Land Warfare, 2425 Wilson Boulevard, Arlington VA 22201, e-mail ncurry@ausa.org or telephone (direct dial) 703-907-2636 or (toll free) 1-800-336-4570, ext. 2636.
# Contents

Preface.................................................................................................................................v
Introduction............................................................................................................................1
The Influence Machine .........................................................................................................2
Manufactured Influence .......................................................................................................3
Linear Strategy in an Exponential World.............................................................................5
Solution Space.......................................................................................................................7
Conclusion.............................................................................................................................10
Endnotes...............................................................................................................................11
Preface

The Defense Department has become fascinated with Artificial Intelligence (AI), and rightly so, as this technology will be as transformative as electricity was a century ago. This paper proposes that the convergence of AI and information operations (IO) represents a greater strategic liability than computationally similar applications—physical adaptations, like drone swarms, and informational applications, such as intelligence process optimization—that have drawn so much budgetary attention. Using a hypothetical “Influence Machine,” the article explores evolving techniques that achieved near-human acuity on many influence tasks. With all the necessary data essentially pre-structured, machine learning can perform these tasks at a massive scale. Using the historical parallel of the impact of the printing press during the Reformation, the author unpacks the exponential potential of emerging influence systems. He also examines how, applied during a time of crisis, such tools may provide a strategic defeat mechanism. In conclusion, he provides two broad recommendations and three specific techniques as examples of the kind of innovation needed to out-disseminate opponents—at scale. Regardless of what course U.S. competitors pursue, Influence Machines will progress and proliferate because the attention trade is highly lucrative. Future IO systems, built by competitors and corporations alike, will be able to simultaneously monitor and affect tens of thousands of people at once, but the Pentagon’s current efforts to integrate AI lack sufficient IO functionality. The pieces are already there, waiting for an adversary to assemble its own Influence Machine, for which disinformation is simply a user setting.
The Influence Machine: Automated Information Operations as a Strategic Defeat Mechanism

The printing press is the greatest weapon in the armory of the modern commander.

T.E. Lawrence

Introduction

The Department of Defense (DoD) is anxious about artificial intelligence (AI); some experts are worried about drone swarms, others excited about machine-aided decisionmaking, and still others are concerned that advanced computing will enable robots to kill without human control. U.S. defense leaders are right to be alarmed. Andrew Ng, a prominent AI developer, believes that “[j]ust as electricity transformed almost everything 100 years ago, today I actually have a hard time thinking of an industry that I don’t think AI will transform in the next several years.” The transformation of one industry in particular has grave implications for U.S. national security: influence. AI-guided information operations (IO) utilize tools that can shape a target audience’s perceptions through the rapid and effective mimicry of human empathy with that audience. Machine speed influence operations are occurring right now, but future IO systems will be able to individually monitor and affect tens of thousands of people at once. Though the threat of automated influence exists quite literally on the smartphone in front of you, the Pentagon’s current efforts to integrate AI do not appear to include any reasonably resourced IO response.

Automating IO decisions of whom to target and when, where and how to affect that audience, enables a line of effort that is more impactful, at a strategic level, than other applications of AI. This paper addresses the threat of automated influence—purposeful use of AI to monitor specific audiences and produce and distribute misleading information to them over digital media networks for the purposes of foreign security objectives. Russia, the pacing threat for the DoD, is leading state development of these tools, as exemplified by their attacks on the 2016 U.S. election; the trail they have blazed is being followed by others. Three particular capabilities enable this new threat vector: algorithmic content generation, personalized targeting and firehose dissemination. For the purposes of this discussion, their theoretical convergence is termed “the Influence Machine.” This term will be used to explore the exponential benefits of AI-driven influence operations and to compare the reach and scale of possible target effects versus current application of AI across DoD, such as physical/robotic and informational mission
command applications of AI. Historic examples, such as the printing press and documented Russian uses, provide precedent for responding to the evolved threat. The paper concludes with an analysis of the Influence Machine’s strategic impact and existing measures that are failing to address an already apparent risk, arguing that the U.S. Government (USG) must not only allocate increased resources to a response, but also enact policies that can compete with the scale of the threat.

The Influence Machine

Amidst the dizzying pace of today’s technological change, it is important to first identify the technologies and techniques most capable of strategic disruption. As shrinking transistors first decreased the cost of simple computation, AI now reduces the cost of prediction, according to economist Ajay Agrawal. More specifically, what online marketing firms and authoritarian governments have done is convert human influence into a prediction problem, an almost infinitely diverse set of if/then statements. Agrawal adds that AI can provide the ability to “take information you have and generate information you don’t have.” In this way, persuasive content can be automatically mass-produced. Adding machine learning to IO allows users to microtarget the audiences most susceptible to the latest behavioral psychology techniques, to exploit emotion and bias and to concentrate on those target groups that are best placed to affect the desired outcome. With that guidance, customized mental munitions can be fired at machine gun speed.

These computer-driven capabilities are not entirely new. For years, algorithmic content generation tools, like those of the company Narrative Science, have helped writers to construct sports stories and stock summaries by using natural language processing to turn structured data streams into readable prose. It is just as easy to use them to create disinformation, such as a widely-shared computer-contrived video, where President Obama appears to provide a warning about deepfakes—doctored images that were once limited to pornography—will challenge our presumptions about truth in the coming years.11 The segment reveals Jordan Peele, a noted Obama impersonator, having mapped the former President’s face onto his own in order to deliver a warning about emerging technology, an entertaining yet frightening demonstration of content manipulation. Regarding deepfakes and social media dissemination, Senator Marco Rubio recently asserted that there is no “individual, political campaign, [or] any organization with bandwidth to knockdown the spread of that information fast enough.” With the increasing prevalence of video content and the appearance of immersive technologies such as augmented reality, computer-generated content will only become more pervasive and less distinguishable from reality.

Still, even if malign actors can lie at machine speed, they still have to get the story to an audience; personalized targeting is required for the Influence Machine to know who will accept a particular message. The more mundane prediction application for AI is foreseeing what a target human will want or do, the way Amazon uses its algorithms to recommend purchases. Gaussian mixture models and/or Naïve Bayes can do this for any company. Programmatic marketing, using consumers’ data habits to drive real-time automated bidding on personalized advertising, has been in use for a few years. Cambridge Analytica’s Facebook targeting, for instance, made international headlines using similar techniques. AI trained with data from users’ social media accounts, economic media interactions and their devices’ positional data can infer even more predictive knowledge of its targets. Emerging tools like the app Replika can very nearly befriend a person, for good or ill. That AI was built to mimic its lead designer’s
deceased boyfriend; having processed all of the texts and emails he had ever sent, it provided contextualized banter about her current life after his death. The design team published the app after seeing how much people opened up to a machine/friend; two million people had downloaded this advanced chatbot by January 2018. The uniqueness of individually lived experiences should decentralize the inputs to an unknowable variability, but the monetization of social science—at the nexus of tech stocks and marketing fees—has rendered many human behaviors quite predictable for recurrent neural networks. Here, the Influence Machine disproves Clausewitz’s supposition that will or morale cannot be classified or counted.

If the Influence Machine knows whom to target and to what they will respond, its next prediction is when to provide catalytic input. The answer is: always, in firehose fashion. Russian bot armies continue doing this very thing, with high-volume and near realtime targeting. The New York Times maintains about a dozen Twitter feeds and produces around 300 tweets a day, but Russia’s Internet Research Agency (IRA) regularly puts out 25,000 tweets in the same 24 hours. The IRA’s bots are really just low-tech curators; they collect, interpret and display desired information to promote the Kremlin’s narratives and build an audience. The real power of this weapon is its secondary effects. When a cognitive munition successfully impacts its target, those victims then fire the meme into their networks at a sometimes unimaginable pace. If “repetition is a key tenet of IO execution,” then this “machine gun” ability to fire information at an audience will, with faux-empathetic precision and custom content, provide the means to change a decisive audience’s very reality.

Next-generation bot armies, powered by simple reinforcement learning, will employ much faster computing techniques and profit from substantially greater network speeds when 5G services are fielded. To take the next step, no breakthrough science is needed; no bureaucratic project office is required. As Neuromancer author William Gibson maintains, “The future is already here—it’s just not evenly distributed.” These pieces are here, only waiting for adversaries to assemble their very own Influence Machines.

**Manufactured Influence**

Today’s strategically disruptive combination of mass-produced media, targeted content, and decentralized dissemination is not without precedent. Johannes Gutenberg introduced Europe to the movable type printing press in the 1440s; commercially, it was a success, but it did not start a revolution. Fast forward to the 16th century. The University of Wittenberg was founded in 1502, the same year that the town’s first printer arrived. A decade later, Martin Luther began to speculate that this printing technology was something special. The Protestant reformers in Europe vastly out-printed their Catholic opponents in the critical years from 1521 to 1526, after the Edict of Worms, which formally denounced Martin Luther and set the stage for far-reaching sectarian conflict. Critically, the majority of the reformers’ works were published in the vernacular; they were also composed to stoke German nationalism. At one point, in 1523, the reformers were printing four times as many German works as were their competitors in the counter-reformation. Printers in at least 13 towns were producing material for Luther. Wittenberg, chief among them, had 12 working presses; there were at least 50 printers known to have produced content for Luther. Because of the decentralized production, rapid distribution and consumable content, the ideas of the Reformation overwhelmed local officials, producing what Richard Cole has termed a fait-accompli. The Wittenberg Monk’s heretical leaflets paved the way to 130 years of bloodshed between Catholics and Protestants, ending with the 1648 Peace of Westphalia and giving rise to the modern, state-centric world order.
Those pamphlets were printed on Gutenberg's invention, technology with world-changing, strategic impact.

The information explosion of the Reformation provides a useful parallel with contemporary information dissemination, not only for its scale, but because it shared the strategic field with the first battle decided by arquebus fire and the emergence of the first professional armies. As the Influence Machine's significance is overshadowed by quadcopters and intelligence software, the strategic significance of Gutenberg's innovation is minimized by commentary on firearms and fortification. The modern proliferation of social media influence clearly resembles the scale of technology diffusion of the Reformation, but with capabilities for content production and dissemination that are multiplied by several orders of magnitude. The disruption of Protestantism helped keep the princes of Europe internally-focused, unable or unwilling to mount major battles of conquest until the Thirty Years War. Similarly, in the contemporary world, America's competitors would like the United States to be internally-focused while they remake the world order to their liking.

The reporting on Russian operations during the 2016 election suggests that the Kremlin generated significant internal disruption for the United States; Facebook accounts operated by Putin's digital agents generated over 60 real protest events. Pages like HeartofTexas and Blacktivist used digital means to trick American citizens into the street to march in support of a foreign political objective. Consequently, the Russians not only achieved strategic disruption, but also discovered a useful operational tool for unconventional warfare; they tapped into the protest potential of a local population without setting foot in the theater. The Russian-backed mobilization of physical protest through social media is not just an update of their decades-old Active Measures; their Cold War efforts affected other nations' policies in ways that are distinct from espionage but were wholly revolutionary. Futurist David Brin noted that, in the 1920s and 1930s, Nazis and Stalinists began taking advantage of radio and public address speakers, inventions enabled by vacuum tubes. By the 1950s, Soviet Russia possessed a centrally-controlled and wire-integrated network of radio and public address systems for internal influence, a system that could reach nearly two thirds of their population almost instantaneously. The volume and ambition of today's Russian IO campaign is far greater because of the capabilities inherent in integrated circuits, the internet and social media.

Russia's chosen mechanism enables it to attack the people of opposing states in a direct and almost instantaneous way that artillery or drones cannot. The revisionist's regional goal is to break the enemy's will and eliminate a population's support for legitimate authorities. For a peer competitor, the aim may not necessarily be as grand as the "decay of liberal democracy." Rather, the Influence Machine is most useful to manipulate a certain value sentiment, of a certain selectorate, for a certain period of time and in support of some specific policy goal, with negative consequences for the opponent. The Russians understand the power of automated influence, and their doctrine has long supported strategic attacks on national will. Today's leading Russian tactician—General Valery Gerasimov—believes that "the role of non-military means of achieving political and strategic goals has grown, and in many cases, they have exceeded the power of force of weapons in their effectiveness."

The ubiquitous nature of electronic platforms provides a direct link, sans geography or security forces, to influence foreign citizens at a massive scale, with feedback—perhaps even through a user’s facial expressions—that provides for the most difficult function of information operations, i.e., measures of effectiveness. With each click on a malign meme, the competitor
gains cookies, traffic data and a piece of network map to drive further operations.\textsuperscript{48} Applied during a time of crisis, such influence could be the difference between prompt response and crippling indecision, just the sort of reaction that “Gray Zone” actions require.\textsuperscript{49} The Russian success at moving adversaries’ citizens to march lays open the Influence Machine’s capabilities, capabilities to which Washington has yet to respond.

**Linear Strategy in an Exponential World**

Though the Pentagon’s strategies recognize the profound potential of AI and call for an increased focus on advanced computing, the application of AI to drones, especially small ones, tends to get the majority of the attention.\textsuperscript{50} Each of the services are exploring concepts using fully-automated swarms or a group of unmanned aerial vehicles (UAVs) driven by AI, like Perdix or Gremlin, to overcome particular adversary capabilities.\textsuperscript{51} Recent Chinese drone shows, featuring thousands of vehicles in a single display, naturally leads to the question of weaponization—unsurprising, given Russia’s tactic of using drones to drop thermite grenades on Ukrainian ammunition depots.\textsuperscript{52} Renowned defense expert David Kilcullen even said that introduction of small drones at the squad level rivals the invention of the machine gun; the Islamic State (ISIS) deployed quadcopters to great alarm in Syria and western Iraq.\textsuperscript{53}

Certainly, the idea of “Quads-for-Squads,” the popular term for the Marines’ teleoperated drone employment program, is an important tactical innovation.\textsuperscript{54} However, even the sum-total production capacity of China—which has cornered the world drone market—cannot manufacture UAVs at a rate rivaling that of machine gun ammunition production. Nor would the number of dead from drone-dropped munitions in Mosul rival the death toll at the Somme. Also, the Chinese record-breaking drone swarm weighed—in total—thousands of pounds; deploying that system would still require the same lines of communication that an ordinary infantry platoon would need.\textsuperscript{55} A drone swarm, as an indirect fire weapon system with a given range and velocity, is also confined by locality. The advertised distance of the flight for the vehicles in that Chinese display is 1 kilometer, a range comparable to that of a modern sniper rifle.\textsuperscript{56} Whether armed with directed energy weapons or powered by hypersonics, robots in a flat world are limited by classical mechanics and are similarly limited in their strategic impact.\textsuperscript{57}

Controlling killer drones through combinatorial optimization problems in fluid multi-agent systems are only one application of AI in military systems.\textsuperscript{58} Situational awareness and decisionmaking are arguably where the DoD is making the most progress, especially in the intelligence community. The National Geospatial Agency is using AI to hunt North Korean missile sites, and the Algorithmic Warfare Cross-Functional Team is running Project Maven for sorting imagery feeds.\textsuperscript{59} The Army’s business practices are also benefiting: the service is trying predictive maintenance monitoring and AI simulations for tactical target recognition, and its Special Operations Command is using AI to optimize personnel management.\textsuperscript{60} The Air Force’s Project Quantum collates cross-domain sensor data for predictive analysis of Programming, Budgeting and Execution problems.\textsuperscript{61} Though terrorist drones seem to present a significant threat and “Siri-for-Soldiers” appears groundbreaking, AI-guided robotics and mission command optimization present only linear change, whereas automated influence has exponential potential.

These modernization programs are admirable but, like physical autonomy, they do not have an exponential effect or reach. The performance gains from optimization-through-software are measured in percentage points, not orders of magnitude. The Influence Machine is
more strategically dangerous than AI-guided robotics because supply chain logistics and locality both limit drones and any other physical weapon. Neither of these AI applications, robotics or process optimization, provide a truly strategic risk to a nation-state on their own. Nonstate actors have used (and will continue to use) small drones to cause havoc, perhaps even in the United States. Indeed, state versus nonstate adaptation of quadcopters is already occurring in Mexico.62 Some nation-state may decide to send a plane full of slaughterbots to drop on the United States.63 However, the former terrorist event is not a defeat mechanism for the United States, and the latter is an act of war. America’s primary competitors have shown that they want to pursue conflict short of war.64 Though the tactical possibilities of battlefield robotics are immense, and AI-enabled decisionmaking may be operationally decisive, neither employment vector can have the strategic impact of automated influence that is generated by the Influence Machine.

The distribution of slaughterbots can occur at about 25 miles per hour—the distribution of memes on the digital battlefield occurs at close to the speed of light.65 The range of the drone swarm is between 500 meters and a few miles, depending on its autonomy level; the range of a meme is the range of the Influence Machine’s sharing network, often global. The drone also has a distinct point of origin. A meme, much like a cyberspace munition, does not necessarily have a point of origin; tracing it back to a firing position can be difficult. However, unlike offensive cyber operations against, for example, a Ukrainian power grid, the operation environment is essentially permissive, and the battlefield complexity is all but eliminated by the pervasiveness of the targeted platform. Also, the targets and techniques are timeless; the human mind does not get software updates or naturally use two-factor authentications.

In the early days of the consumer internet, one group quickly grasped the high return on investment and almost instantaneous supply chain that digital influence could provide: Nigerian email scammers. These purveyors of one of the oldest social “hacks” still send droves of emails and, though now receiving relatively few replies, they still annually gross upwards of $360 million. The marginal costs to create phishing material, provide VPN/proxy services and maintain Chinese bank accounts is quite low when compared to the reward they reap.66 The Russian state benefits from this same asymmetry every time the IRA tweets. Dmitry Kiselyev, the director of the government-controlled news agency Rossiya Segodnya, maintains that, “today, it is much [costlier] to kill one enemy soldier than during World War II, World War I, or in the Middle Ages; if you can persuade a person, you don’t need to kill him.”67 The Russians adapted an asymmetry that is more threatening than any robot or spreadsheet.

The global and ubiquitous nature of electronic platforms, from social media to email chains, provides an attack surface with exponential potential. Noted futurist and inventor Ray Kurzweil describes this sort of potential as when a “key measurement, such as computational power, is multiplied by a constant factor for each unit of time (e.g., doubling every year) rather than just being added to incrementally.”68 If one can harness such exponential change, they can benefit from the asymmetric gains that Kurzweil describes in his Law of Accelerating Returns, where the benefit from investing energy constantly increases rather than eventually diminishing as expected in normal economics.69 For the Influence Machine, the measurement is people reached, and the constant factor is the rate of viral transmission of a particular meme.70 With any physical weapon technology, save thermonuclear devices, the people reached—killed—moves on some linear scale; for the Machine, the conditions are quite different. On a given day, there are up to 3.9 billion people online, all theoretically within range of a meme; no physical system can possibly affect as many people as the Influence Machine.71 This is not the first
attempt to penetrate deep into an opponent’s population: strategic bombing theory also sought
to reach a campaign decision by coercing the civilian population. Like strategic bombing of
generations past, the Influence Machine aims at massive strikes deep into the state, intending to
attrite the will of the people; but unlike strategic bombing, the destructive event does not create
a shared experience. Instead, the goal is to divide at a personal or tribal level, thereby denying
any value to the target’s collective strategic goals.

The crux of the Influence Machine’s value is the inherent vulnerability of Western democra-
cy, that decisionmakers are beholden to a malleable selectorate. As Senator Mark Warner noted,
“We’re increasingly in a world where cyber vulnerability, misinformation and disinformation
may be the tools of conflict.” By affecting the cognition—the will—of enough people, this
machine can prevent or delay a democratic government’s physical response to aggression; it is
a defeat mechanism. The Influence Machine’s objective comes down to changing the value
of the target’s strategic goal. Clausewitz knew that the political object, the original motive, in
a conflict was the essential factor in any deterrence equation. The smaller the value demanded
of an opponent, the less that competitor would be willing to try to deny it. This is the inverse
of Fearon’s “tying hands” findings that the increase of the perceived costs for an audience, a
national population, tends to prevent a country from backing down when attempting to coerce
an opponent. With automated influence, that opponent attempts to lower the expected benefits,
on the part of the competitor’s audiences, for the intervention action. The Influence Machine
enables defeat before any shots are ever fired by removing “the physical means or the will to
fight.” In this condition, a defeated state’s executive is unwilling or unable to respond to a
threat action, thereby yielding to the opponent’s will. As fake news becomes frighteningly
competitive with real news, the emergence of the Influence Machine presents a novel way to
“hack” the unchanging human nature of war.

Solution Space

This threat adaptation appears at an opportune but chaotic time; the DoD is coming to
grips with concepts like conflict short of armed conflict, information as a joint function and
AI itself. Gravitation toward drone strategies is reminiscent of an age of “thermodynamic
warfare,” in which the ballistic energy needs of ever more physically destructive weapons drive
the economic mobilization for total war. Attraction to machine learning tools that optimize
military informational processes is evocative of “cybernetic warfare” driven by the fantasy
of omniscient knowledge of the battlefield, a dream that should have ended in the jungles of
Vietnam. As part of its adaptation for an age of “chaoplexity”—Antonie Bousquet’s term for
a future characterized by positive feedback loops and non-linearity—the USG must confront
the cognitive influence potential of these technologies in three key ways: funding at a level
commensurate with the threat, designating an authoritative coordination body and adopting
innovative techniques for the scale of the problem.

There are quite a few disparate efforts underway to identify and grapple with the threat of
automated influence. Oxford University’s Computational Propaganda project has cataloged au-
tomated influence in several countries; the German Marshall Fund has created the Hamilton68
database to track Russian influence on Twitter. There are software solutions for spotting deep-
fakes and bots. The Defense Advanced Research Projects Agency—DARPA—toyed with
automating the characterization, but their project only reached 40 percent effectiveness. Even
that meager gain is offset by malign actors’ constant innovations that make them increasingly
difficult to spot. Users can detect bots with low-tech criteria, but those rules only loosely
impede the threat. Efforts like StopFake and Bellingcat have tried to disprove Russian fake news, but perpetrators can always lie faster than any debunking process can operate, and the exposure actually increases the virality of the lie. Rebutting alone will not work: the British Army’s studies on Russia’s propaganda war indicate that responses to misinformation take hours or days, making contradiction almost completely ineffective.

Some have offered institution-level solutions. Brookings’ Polykova and Boyer have identified that “technological advances in AI, automation and machine-learning, combined with the growing availability of big data, have set the stage for a new era of sophisticated, inexpensive, and highly impactful political warfare.” However, their reasonable “whole-of-society” recommendations—such as government, private and non-profit investment in software to identify computational propaganda or improvements to information sharing—are not exponentially potent enough to meet the challenge posed by automated influence. A recent RAND report on the firehose of Russian falsehood offers recommendations such as directing information streams at the competitor’s propaganda targets rather than the originator himself, as well as increasing the flow of native persuasive information. Publicly, American digital IO has been quite limited. Confined to military social media “WebOps” and State Department Global Engagement Center (GEC) actions, these efforts have been contractor-heavy, prone to naïve missteps and fixated on counterterrorism. Some dabbling in AI support to IO is occurring in the area of social media analysis. The USG must do more than just monitor the situation; it must produce actual effects in the environment.

The threat of automated influence demands solutions that can hack the system faster than rapid-fire disinformation and not merely respond to it. The Russian adaptations to the opportunities of automated influence are already at a scale that far surpasses the capability and scale of one-off websites and a few social media pages. The Kremlin has perhaps thousands of state employees and contractors ready to inject malign content into the local environment. According to a recent RAND report, the Kremlin uses a high number of channels and a dizzying array of messages to overwhelm the audience with rapid, continuous and repetitive exposure. When this Russian firehose turns on, the stickers and posters that fueled the last generation of political arguments will not be enough to preserve Western cognitive integrity. The United States still possesses the most formidable military in human history, but that may not matter if an aggressor is able to neutralize the will to employ it. Leaders may be tempted to look away from the problem because the Pentagon lacks the authority, and even the mandate, to contend with the security of the domestic cognitive environment. However, as China invests heavily in AI and Russia continues its influence campaign, America must respond.

The DoD will continue to grapple with the ethics of automated killing, with the terrorist threat of small drones, with the “black box” problem of AI support to mission command—and with Google employees not wanting to help. Of all of these issues, automated influence operations remain the most pressing strategic threat among the more lethal or tangible AI employment venues. Unfortunately, DoD’s current outlays for integrating AI do not appear to include the necessary resources to adapt. The Pentagon is spending almost $7 billion a year for robotics and practically none on automated influence, while China operates a $1.3 billion state media apparatus. The USG may not need a Manhattan Project to meet the challenge of automated influence; neither the Russians nor Martin Luther needed a crash program for their innovations. The 2019 National Defense Authorization Act already requires a study on AI topics; as a bare minimum, automated influence must be included. Further, the new Joint Artificial Intelligence Center and its $1.7 billion budget allocation must have an IO function.
The USG poured billions of dollars into the counternarrative fight against jihadi extremism.\textsuperscript{97} As America begins to refocus on great power competition, the response to AI-driven IO must be appropriately resourced.

Notably, the USG does have some money committed to digital IO, with over $100 million for the GEC and upwards of $500 million for WebOps at the Combatant Commands.\textsuperscript{98} Unfortunately, even if some official decides that AI-driven IO is an issue, there is no single organization to integrate and synchronize U.S. adaptation to the threat of the Influence Machine, much less to supervise the use of any such machines that America might build. A whole-of-government integration function is needed as a foundation to any response. The Russian government has integrated a comprehensive concept for maneuvering within a “unified information space” with the General Staff as the coordinating authority, the focal point being the National Centre for Direction of the Defence of the Russian Federation.\textsuperscript{99} Many have argued to restore a U.S. Information Agency. Though the USG does not need relics, it does need something to run its operations. The potential political backlash, like that encountered by Secretary Rumsfeld’s Office of Strategic Influence, must simply be stomached and a coordinating authority for IO, sometimes unattributed and even provocative, must be established in earnest.\textsuperscript{100} And, while the DoD might be a trusted actor and at risk from automated influence, the military cannot sit at the head of this effort—though it must be able to engage, to shoot back, at a rate that is competitive with America’s adversaries.

The 2016 Countering Disinformation and Propaganda Act casts DoD in a supporting role against the threat of automated influence of domestic audiences, but the Department desperately requires techniques to out-disseminate its opponents among indigenous constituencies without whose support victory would be impossible or extremely costly.\textsuperscript{101} Competitors must be beaten at scale, with truthful content. Given the current state of technology, the DoD has the necessary data to create a better, more truthful firehose. \textit{The U.S. military must curate faster than its opponents can lie}, by filtering streaming field footage, creating content out of existing mission command feeds and aiding public affairs’ functions with chat-bots. The DoD needs radical transparency, though certainly with a selective eye; existing AI capabilities can provide it. The viral nature of ISIS combat footage on social media—the Chechens were arguably the first to upload viral combat brutality—and the outcry generated by U.S. combat footage captured in Niger indicate that such content is inherently powerful.\textsuperscript{102} There is a duopolistic market, with significant first-mover advantage, for information coming from the frontlines. U.S. Soldiers already carry cameras on patrol and are beginning to use drone cameras as the fighters in Syria and Ukraine do. The DoD must develop ways to exploit that sort of data. AI image recognition tools provide the means to sort and prioritize footage coming from any number of camera-equipped Soldiers. This combat footage could be a powerful and effective tool, but is instead being wasted.

It is also worth noting that software such as Narrative Science can produce press releases and social media updates with little to no man-hour investment. The U.S. military network produces terrabytes of mission command data on scalar movement, system state and contextual position. These numerical data streams fit the same criteria as sports statistics, stock prices and election polling data that have allowed journalists to become part cyborg.\textsuperscript{103} The same software that is replacing humans for discovery in court cases can automate part of the military’s media relations function; no human Public Affairs officer can be expected to have memorized every line of every DoD policy document that a reporter might grill them on, but a machine can.\textsuperscript{104} The natural aversion to machine-driven transparency is operations security, revealing details
that might compromise a mission, but if Wall Street can learn to trust tens of billions of dollars in trades to AI, perhaps the military can let a computer manage a few social media accounts.\textsuperscript{105} Each of these three techniques is in step with Google’s AI Code of Conduct; none needs to have its development cloaked in secrecy.\textsuperscript{106}

**Conclusion**

AI, as a concept, has been around for decades; for most of that time it did not live up to expectations. Now, even the stoic Secretary Mattis, confronted with the power of this contemporary technology, has questioned his confidence that the fundamental nature of war will not change.\textsuperscript{107} His organization must also readdress its base assumptions about influence and AI. While the Marine Corps shops for drone swarms, the Navy for electronic warfare enablers, the Air Force for resilient satellite networks and the Army for manning-optional vehicles, computational propaganda has proliferated to at least 12 countries on five continents.\textsuperscript{108} Machine guns and drone swarms can destroy or disrupt, but they cannot defeat another nation. One must ask, what had greater impact on U.S. strategic options in the last half of the 20th century: the antitank guided missile, Airborne Warning and Control Systems or the advent of 24-hour cable news? The enemy could not directly manipulate CNN, but the Influence Machine allows them to intentionally affect today’s digital media. And, major news networks increasingly rely on social media, rather than expensive string reporters, for information, as evidenced by Fox News’ use of the Russian @TEN_GOP account as a source.\textsuperscript{109}

With lower marginal cost, greater range, higher production rate and the potential for exponential effects, the Influence Machine poses a bigger threat and presents a greater opportunity than other disruptive uses of AI. For these reasons, the technology that powers the Influence Machine will continue to progress, regardless of competitor nation behaviors or USG response. Manufacturing content, true or not, and disseminating it to the most valuable audience is big business; this application development market has turned *The Washington Post* from a news outlet into a software company.\textsuperscript{110} Though narrative creativity is still a nascent venture for bots, IBM and Google also have computing projects whose applications make reasoned arguments that mimic human conversation.\textsuperscript{111} Private industry will continue to refine the automatic weapons of influence described above; commercial attempts to profit from the shortcomings of human cognition with ever-greater computing power and connectedness will drive this arms race.\textsuperscript{112}

In contrast to the ideal future, where the internet was hoped to be a public good that featured protected speech and freer audiences, antidemocratic forces have taken advantage of connectedness to create distraction and doubt. America’s competitors have already exploited the cognitive conflict space created by the monetization and digitization of the human experience. Though foreseen decades ago, this strategic disruption has yet to receive an effective response and has now begun to infest the cognitive spaces of home-town America.\textsuperscript{113} Though automated influence cannot necessarily win a campaign, it can preclude democracies from fighting one.\textsuperscript{114} The convergence of automated influence tools driven by developments in AI represents a strategic liability of greater consequence, as a defeat mechanism, than the computationally similar physical robotic systems or informational optimization techniques that have, thus far, drawn so much budgetary attention. The U.S. national security enterprise requires AI-driven influence policy, technology integration and significant numbers of innovative techniques that can generate effects as virally as threat influence systems. The problem set is in no way unprecedented, but exponential acceleration of this technology increases the opportunity cost of inaction at the same rate.
Endnotes


5 This definition borrows heavily from the definition of “Computational Propaganda” from the Oxford project of the same name; the main difference here is that content generation and security objectives are included as an end.


8 Ajay Agrawal, “A.I. economics: How cheaper predictions will change the world,” Big Think (4 May 2018), https://www.youtube.com/watch?v=YRzGSp_bO1M.

9 Ibid.


23 Chui et al., “An Executive’s Guide to A.I.”


26 This count of The New York Times inputs is the author’s, from December 2017; The IRA count is from the same period and sourced from the Alliance for Securing Democracy, Hamilton 68: Tracking Russian Influence Operation on Twitter (Washington, DC: German Marshall Fund, October 2017), http://dashboard.securingdemocracy.org.


29 Chui et al., “An Executive’s Guide to AI.”


32 Ibid.


34 Ibid.


38 Howard, War in European History, 27.


42 David Brin, Naval Postgraduate School, 7 July 2017.


Watts, Messing with the Enemy, 195.

The “Gray Zone” as formalized by General Votel, et al., is conceptual space “where threats and our response to those threats will take place in a segment of the conflict continuum that . . . is characterized by intense political, economic, informational, and military competition more fervent in nature than normal steady-state diplomacy, yet short of conventional war,” Joseph L. Votel, Charles T. Cleveland, Charles T. Connett and Will Irwin, “Unconventional Warfare in the Gray Zone” Joint Forces Quarterly 80, no. 1 (2016): 101.


@CHACR_Camberley, “‘The introduction of small drones for vertical capabilities at the squad level rivals the importance of the introduction of the light machine gun’ Dr. Kilcullen. Is it time we approached low level tactics differently? #RUSILWC,” Centre for Historical Analysis and Conflict Research (19 June 2018), https://twitter.com/CHACR_Camberley/status/1009006104695791618.


This is based on the Ehang Ghostdrone, with a weight of 2.4 pounds, that was used in China’s display, https://www.digitaltrends.com/cool-tech/ehang-drone-display-world-record/.


Ibid.

The use of the term “meme” in this paper, originally from Richard Dawkins, refers to any sort of sharable image or post “based on a common theme that has spread widely on the internet” which is also a “vehicle for political messages, used to spread aggressive or racist messages and to incite hatred”; Emerging Technology from the arXiv, “This is where internet memes come from” MIT Technology Review (11 June 2018), https://www.technologyreview.com/s/611332/this-is-where-internet-memes-come-from/.


Automated influence is a method through which the competitor has chosen to accomplish his mission against enemy opposition, therefore it is the “mechanism” defined by ADP 1-02, *Operational Terms and Military Symbols* (Department of the Army: Washington, DC, 16 November 2016).


ADP 1-02, 1-26.

Ibid.


Antoine Bousquet uses “thermodynamic” and “cybernetic” to describe the changing character of contemporary war in *The Scientific Way of Warfare: Order and Chaos on the Battlefields of Modernity* (New York: Columbia University Press, 2009), 32.

Ibid., 34.

Ibid., 30.


Will Knight, “How to tell if you’re talking to a bot.”


Paul and Matthews, “The Russian ‘Firehose of Falsehood’ Propaganda Model.”


Paul and Matthews, “The Russian ‘Firehose of Falsehood’ Propaganda Model.”

Ibid.


Lasconjarias and Larsen, “Introduction: A New Way of Warfare.”


Watts, Messing with the Enemy, 162, 165.


Knight, “How to tell if you’re talking to a bot.”

