DAKOTA DIGITAL REVIEW

AI'S ENERGY APPETITE

▪ *Mark P. Mills*

NORTH DAKOTA'S POLYTECHNIC INSTITUTION

▪ *Douglas Jensen*

THE FITBIT MURDER

▪ *Arica Kulm*

DAKOTA DIGITAL ACADEMY
NORTH DAKOTA UNIVERSITY SYSTEM

FALL/WINTER 2023-24

Introduction to the DAKOTA DIGITAL ACADEMY

KENDALL E. NYGARD, PHD

Director, Dakota Digital Academy, North Dakota University System
Emeritus Professor, Department of Computer Science,
North Dakota State University
Contact: kendall.nygard@ndus.edu

Digital technologies are transforming society and driving revolutionary changes in the world of work. In response, the Dakota Digital Academy (DDA) was founded by the North Dakota University System in the fall of 2020 to provide online education in computing and the cyber sciences. DDA serves students at higher education institutions across the state—as well as residents in the workforce seeking to upskill or change careers—by imparting relevant digital skills. To date, we have focused primarily on courses and certificate programs in cybersecurity and software development.

DDA works cooperatively with the state's 11 public colleges and universities, which include two research universities, four regional universities and five colleges. Also affiliated are North Dakota's five tribal colleges. Talented faculty across the state system work together to design and deliver location-agnostic workshops, full courses, short skill-specific courses and certificate programs. Some activities are oriented toward improving the skills of technical people already in the workforce. Others focus on continuing education and credentialing for K-12 teachers.

Also included in DDA's instruction are soft skills related to the liberal arts, such as teamwork, creative and critical thinking, problem-solving, ethics and communication, along with considerations of technology's social implications.

Over the last three years, DDA successfully launched Dakota Digital Review, Dakota Digital Discussions and the Workforce Advisory Council, which is comprised of business, industry and government leaders who support DDA's workforce readiness and cyber-educational mission.

Going forward, DDA is pursuing several highly relevant initiatives. One focuses on digital literacy in general education across the university system. Gen Ed refers to suites of required courses imparting knowledge and skills fundamental to all major fields of study and to success after graduation. Increasingly today, literacy in computing and cyber sciences constitute essential components of every student's formation.

A second initiative concerns advancing education in artificial intelligence and machine learning—including content creation systems such as ChatGPT, which are revolutionizing and disrupting nearly every industry, and augmenting or supplanting functionalities that involve reasoning, perception and creativity, which have been strictly human domains throughout history. ▣

■ *Dakota Digital Discussions Webinar Series*

Dakota Digital Discussions is a webinar series presented in the fall and spring semesters by Dakota Digital Academy and Dakota Digital Review. The webinars focus on the digital transformation of our economy, military and society, as well as digitization's profound ethical, legal, cultural, educational and policy implications, including impacts on the arts and humanities and the human psyche. Most Dakota Digital Discussions engage for an hour and are scheduled at noon Central.

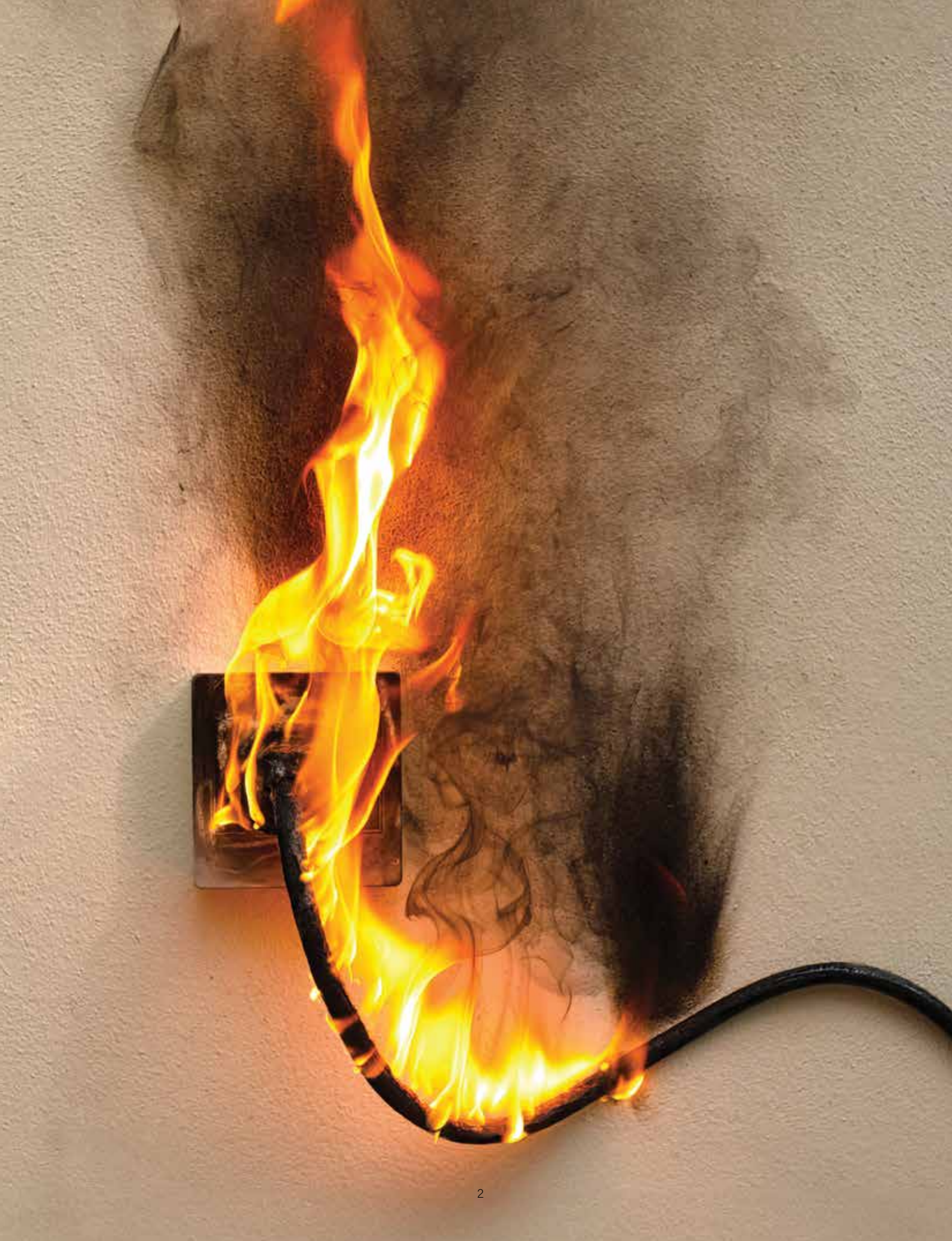
Please access upcoming and archived webinars at: https://dda.ndus.edu/ddd-overview/

DAKOTA DIGITAL REVIEW

- **Dakota Digital Review** publishes articles, in print and online, about digitization and related technologies, as well as digitization's profound implications for our culture, economy, military, political institutions and policies, legal frameworks, moral foundations, and the arts and humanities.
- **Dakota Digital Review** is non-partisan, dedicated to free speech and academic freedom. The articles are written by subject experts in business, industry, government and academia for the general educated reader.
- **We aim to better prepare** the polity and populous to make critical decisions about our collective future and about our individual and family lives.
- Please contact Patrick J McCloskey, Editor, Dakota Digital Review, for submissions, comments or questions: **patrick.mccloskey.1@ndus.edu**.

CONTENTS ▪ FALL/WINTER 2023-24 ▪ dda.ndus.edu/ddreview/

Design / Jerry Anderson. Cover illustration / Tom Marple.



AI'S ENERGY APPETITE: Voracious & Efficient

Mark P. Mills, Senior Fellow, Manhattan Institute

The invention of useful artificial intelligence (AI), epitomized by the hype over ChatGPT, is the latest example of a basic truth about technology. There have always been many more inventions that use energy than those that can produce it. Such is the nature of progress in all domains from medicine and entertainment to information and transportation. While there's a lot of debate, and angst, about AI's implications for the economy, jobs and even politics, there's no debate about the fact that it is a big deal and applications for using it are growing at a blistering pace. It's obvious, but worth stating, that the invention of the car, for example, also 'invented' demand for energy to build and operate cars. Now, in the still short history of silicon chips—the engines of the information age—the arrival of AI promises an unprecedented boost to future energy demand.

Consider an analogy. Imagine it's 1925 and there was a global conference that convened internal combustion engineers.

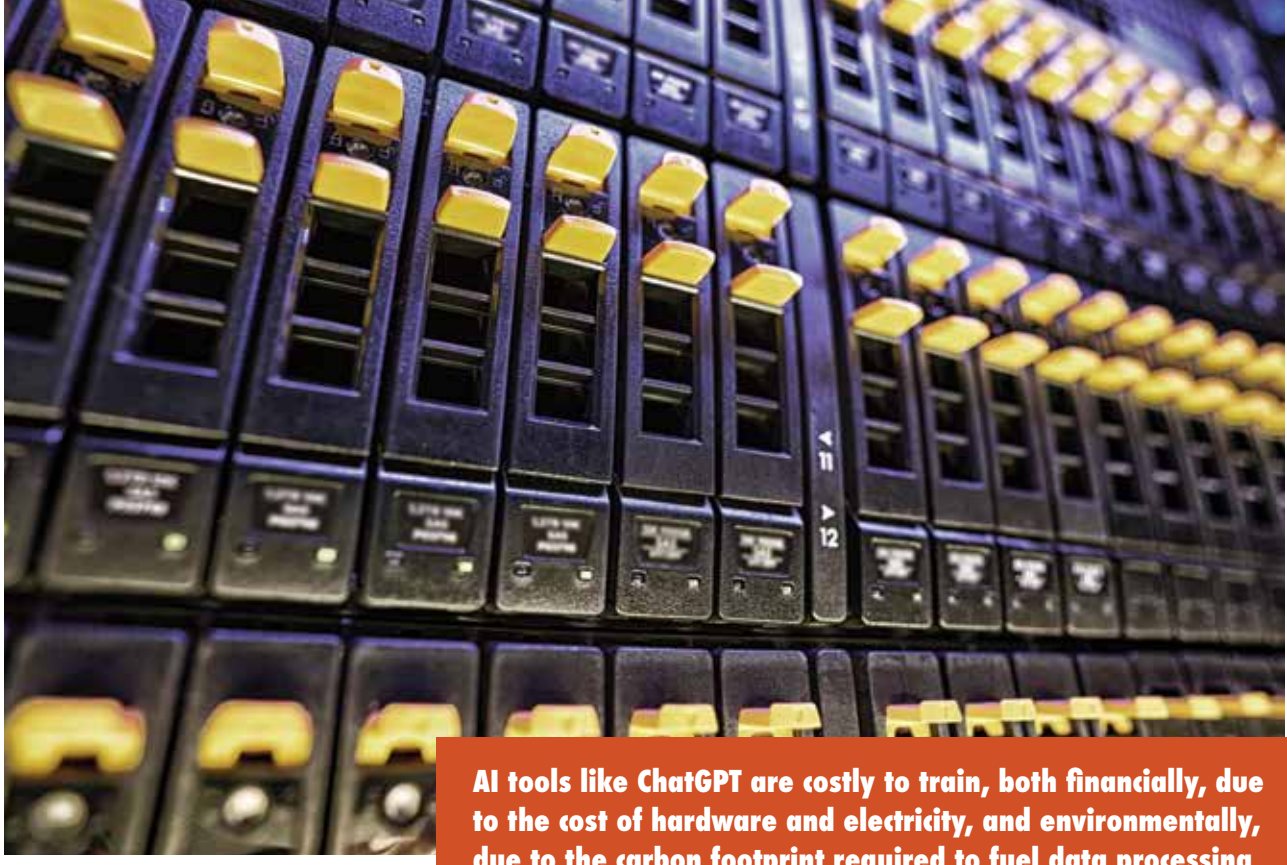
At that point in history, combustion engines and cars had been around for almost four decades, and engines had become good enough in the previous decade to enable building aircraft. Also imagine a Ford engineer at that conference forecasting engine progress that would soon make it practical for widespread air travel. (Yes, a Ford engineer because that company pioneered not only aviation engines, but one of the first practical passenger aircraft, the 1925 TriMotor.[i]" And imagine if that engineer had suggested that global air traffic would become so common that aviation in the foreseeable future would consume nearly as much energy as the entire U.S. used in 1925 for all purposes. A year later, 1926, saw the launch of America's first regularly scheduled, year-round commercial passenger service. The rest, as they say, is history.

Fast forward to an actual conference in 2022—one that took place several months before the November unveiling of ChatGPT—where the CTO of AMD, a world-leading AI chipmaker, talked about how powerful AI engines have become and how fast they were getting adopted. He also showed a graph forecasting that by 2040, AI would consume roughly as much energy as the U.S. does today for all purposes.[ii] If those trends continued, AI would end up gobbling up most of the world's energy supplies. Of course, the trends won't continue that way, but that doesn't change the fact that, as another engineer at AI chipmaker ARM said at that conference: "The compute demand of [AI] neural networks is insatiable." And the growth of AI is still in the early days, equivalent to aviation circa 1925 or, in computer-history terms, to the pre-desktop 1980s era of mainframes.

Inference as Efficient Energy Hog

It's not news to the computer engineering community that AI has a voracious energy appetite. AI-driven "inference," rather than conventional "calculation," is the most power-intensive use of silicon yet created.[iii] That reality is starting to leak out because of the popularity of today's first-generation AI. We see headlines about AI having a "booming … carbon footprint"[iv] and that it "guzzles energy."[v]

Consider, for example, the results of one analysis of what it takes to just build, never mind operate, a modest AI tool for one application, i.e., the equivalent of the energy used to build an aircraft, not fly it. The analysis found that the training phase—training is how an AI tool is built—consumed more electricity than the average home does in 10 years.[vi] A different analysis looked at what it likely took to build a bigger AI tool, say to train ChatGPT, finding that it used as much electricity as an average home in a century.[vii] Of course, like automobiles and aircraft, energy is also used to operate AI, something engineers call "inference," to, say, recognize an image or an object, or give advice or make a decision, etc. Energy use by inference is likely tenfold or more greater than for training.[viii] As one researcher put it, "it's going to be bananas."



AI tools like ChatGPT are costly to train, both financially, due to the cost of hardware and electricity, and environmentally, due to the carbon footprint required to fuel data processing hardware.

AI as General Purpose Tech

There are few intersections of the world of bits and atoms—of software and hardware—that so dramatically epitomize the inescapable realities of the physics of energy, and the challenges in guessing future behaviors and thus energy demands. No one can guess the number of important, or trivial, things where we will want to use AI. And, in fact, for most people in the (lucky) wealthy nations, most energy is used for things other than mere survival. Aviation, for example, is dominated by citizens traveling for fun, vacations, to see family; business travelers account for well under 20 percent of global air-passenger miles (even less now, during the recovery period from global lockdowns). However, unlike aircraft, which are specialized machines used to move goods and people, AI is a universal tool, a "general purpose technology" in the language of economists, and thus has potential applications in everything, everywhere. It's far harder to forecast uses of general-purpose technologies.

AI tools will be put to work for much more than just fine-turning advertising, or performing social media tricks, or creating "deepfakes" to spoof hapless citizens, or making self-driving cars (eventually) possible. AI's power and promise, as its practitioners know, are leading to the potential for such applications as hyper-realistic vehicle crash-testing, or monitoring and planning ground and air-traffic flows, or truly useful weather forecasting. The most profound applications are to be found in basic discovery wherein AI-infused supercomputers plumb the depths of nature and simulate molecular biology "*in silico*" (an actual term), instead of in humans in order to both accelerate discovery and even, eventually, to test drugs. The number and nature of potential applications for AI is essentially unlimited.

Datacenters & Infrastructure Buildout

Analysts have pointed out that the compute power—and derivatively energy—devoted to machine learning has been doubling every several months.[ix] Last year, Facebook noted that AI was a key driver causing a one-year doubling in its datacenter power use.[x] And this year, Microsoft reported a 34 percent "spike" in water used to cool its datacenters, an indirect even if unstated measure of energy use.[xi] It's the equivalent of measuring the flow of water to cool a combustion

The world today has over a billion miles of information highways comprised of glass cables along with four million cell towers that forge an invisible, virtual highway system that is effectively another 100 billion miles long.

engine rather than counting gallons of fuel burned. The need for cooling comes from the energy use.

It is an open secret that AI will drive a massive infrastructure buildout. As a Google VP observed, deployment of AI is "really a phase change in terms of how we look at infrastructure." One article's headline captured the reality: "The AI Boom Is here. The Cloud May Not Be Ready."[xii] Every computer vendor, chipmaker, software maker and IT service provider is adding or expanding offerings that entail AI. It is a silicon gold rush last matched in enthusiasm and velocity during the great disruption in information systems of the 1990s with the acceleration from mainframes to desktops and handhelds. AI enthusiasm is also seen in the stock market where, odds are, history will repeat as well: a boom, a bust and then the long boom. It's the coming long boom that has implications for forecasting energy demands from AI.

To continue with our analogy, while infrastructure growth points to potential for future energy use, it doesn't predict actual outcomes any more than counting highway or runway miles is predictive of fuel use, except that the infrastructure is what enables the fuel to be used. Future historians will see today's Cloud infrastructure as analogous to the 1920s stage of transportation infrastructure in the pre-superhighway days, also a time of grass runways. Even before we see what the next phase of silicon evolution brings—the kinds of services and social changes that will echo those brought by the automobile and aviation—it's possible to have some idea of the scale of energy demand that will bring by considering the current state of today's information infrastructure.

Unlike automobiles and aircraft though, the energy used by digital engines is hidden from plain sight inside thousands of nondescript warehouse-like datacenters. There we find, in each one of them, thousands of refrigerator-sized racks of silicon machines. Each such rack burns more electricity annually than 15 homes, that's before the racks are filled with AI silicon.

Datacenters, in square footage terms, are the skyscrapers of the modern era, except that there are far more of the former and many more being built. And each square foot of datacenter uses 100 times the

power of a square foot of skyscraper—again, before the infusion of AI.

Datacenters of course are only useful if connected to people and other machines and vice versa. The world today has over a billion miles of information highways comprised of glass cables along with four million cell towers that forge an invisible, virtual highway system that is effectively another 100 billion miles long.[xiii] Machinery for transporting bits uses energy just as it does in the world of atoms. And the convenience of wireless networks comes with an energy cost up to 10 times more energy per byte,[xiv] not unlike a similar principle of physics that makes flying more fuel-intensive than driving. Flight-shamers might take note: In energy-equivalent terms, today's global digital infrastructure already uses roughly 3 billion barrels of oil annually, rivaling the energy used by global aviation. And that number is based on data that is a half-dozen years old. Since then, there's been a dramatic acceleration in datacenter spending[xv] on hardware[xvi] and buildings[xvii] along with a huge jump[xviii] in the power density of that hardware, all of that, again, before the acceleration as AI is added to that infrastructure. A single, simple AI-driven query on the Internet can entail over fourfold the energy use of a conventional query.[xix]

It's not that digital firms are energy wastrels. In fact, silicon engineers have achieved epic, exponential gains in efficiency. But overall demand for logic has grown at an even faster, blistering pace. You can "take to the bank" that history will repeat here too: demand for AI services will grow faster than improvements in AI energy efficiency. The cloud is already the world's biggest infrastructure and seeing it expand yet by several-fold, or more, would be entirely in keeping with historical precedent.

Jevons Paradox

Nonetheless, the forecasters and pundits who are preoccupied with reducing society's energy appetite always offer energy efficiency[xx] as a "solution" to the energy "problem." They have it backwards. Efficiency gains have always been the engine that drives a growth, not a decrease, in overall energy use. It's a feature, not a bug, in technology progress, and one that is most especially true in digital domains. This seeming contradiction has been called Jevons Paradox after the British economist William Stanley Jevons who first codified the economic phenomenon of efficiency in a seminal paper published back in 1865. That paper was focused on the claim, at that time, that England would run out of coal given the demands for that fuel coming from a growing economy, growth that itself was caused by the fuel of industrialization. The solution offered by experts at that time was to make coal engines more efficient.

Jevons, however, pointed out that improvements in engine efficiency—i.e., using less coal per unit of output—would cause more, not less, overall coal consumption. Thus, the ostensible paradox: "It is wholly a confusion of ideas to suppose that the [efficient] use of fuel is equivalent to a diminished consumption …. new modes of [efficiency] will lead to an increase of consumption."[xxi] Some modern economists call this the "rebound effect."[xxii] It's not a rebound as much as it's the purpose of efficiency.

Put differently: the purpose of improved efficiency in the real world, as opposed to the policy world, is to make it possible for the benefits from a machine or a process to become cheaper and available to more people. For nearly all things for all of history, rising demand for the energy-enabled services outstrips the efficiency gains. The result has been a net gain in consumption.

If affordable steam engines had remained as inefficient as when first invented, they would never have proliferated, nor would the attendant economic gains and associated rise in coal demand have happened. The same is true for modern combustion engines. Today's aircraft, for example, are three times more energy efficient than the first commercial passenger jets. That efficiency didn't "save" fuel but instead propelled a four-fold rise in aviation energy use.[xxiii]

The same dynamic is at play with today's digital engines, the driving force of the 21st-century economy. In fact, the microprocessor represents the purest example of the Jevon's paradox. Over the past 60 years, the energy efficiency of logic engines has improved by over *one billion* fold.[xxiv] Nothing close to such gains

are possible with mechanical and energy machines that occupy the world of atoms.

Consider the implications from 1980, the Apple II era. A single iPhone at 1980 energy-efficiency would require as much power as a Manhattan office building. Similarly, a single datacenter at *circa* 1980 efficiency would require as much power as the entire U.S. electrical grid. But *because* of efficiency gains, the world today has billions of smartphones and thousands of datacenters. We can only hope and dream that the efficiency of AI engines advances similarly.

A leading-edge AI chip today delivers more image processing capability than a supercomputer could just two decades ago. In a sign of our times, last year the silicon start-up Cerebras introduced a kind of Godzilla-class AI chip the size of an entire silicon wafer (think, medium pizza) with more than two trillion transistors and a 15-kilowatt power appetite. That's more peak power than used by three houses. But it offers more than a ten-fold gain in efficiency over the best AI chips. Competitors will follow. That's why the market for AI chips is forecast[xxv] to dominate semiconductor growth and explode some 700 percent in the next five years alone.

Green vs AI

Of course, the Jevons paradox breaks down in a microeconomic sense, and for specific products or services. Demand and growth can saturate in a (wealthy) society when limits are hit for specific items regardless of gains in efficiency, e.g., the amount of food a person can eat, or the miles-per-day one is willing to spend driving, or the number of refrigerators or light bulbs per household, etc. But for such things, we're a long way from saturation for over two-thirds of the world's citizens. Billions of people in the world have yet to become wealthy enough to buy even their first car or air conditioner, never mind use an AI-infused product or service.

But one can understand why the "green AI" community is alarmed over what will come. Even before the Age of AI is in full swing, today's digital infrastructure already uses twice as much electricity as the entire country of Japan. We await the Cloud forecasts that incorporate the energy impact of the AI gold rush. As Deep Jariwala, a professor of electrical engineering at the University of Pennsylvania provocatively put it: "By now, it should be clear that we have an 800-pound gorilla in the room; our computers and other devices are becoming insatiable energy beasts that we continue to feed."[xxvi] Prof. Jariwala went on to caution: "That's not to say AI and advancing it needs to stop because it's incredibly useful for important applications like accelerating the discovery of therapeutics." There's little to no risk that governments will or can directly throttle AI development (though some have proposed as much). Ironically though, government energy policies could make AI expensive enough to slow deployment.

Consider a simple arithmetical reality: In a low-cost state, training a high-end AI requires buying about $100,000 in electricity, but you'd spend over $400,000 in California. And the training phase for many applications is necessarily repeated as new data and information are accumulated. One can imagine, as some have proposed, doing the training at remote locations where electricity is cheap. That's the equivalent of, say, buying energy-intensive aluminum to build airplanes from places where energy is cheap (China, using its coal-fired grid, produces 60 percent[xxvii] of the world's aluminum). But most uses for AI require operating it and fueling it locally and in real-time, much as is the case to operate an airplane. Wealthy people will be able to afford the benefits from higher cost AI services, but that trend will only further widen the "digital divide," wherein lower income households are increasingly left behind.

Transition Irony

There is some irony in fact that many in the tech community have joined with the "energy transition" lobby to promote the expansion of power plants that are not only increasing the cost of electricity but making it more difficult to deliver it when markets need it. Despite the mantra that wind and solar are cheaper than conventional power plants, the data show that, in every state and every country, the deployment of more episodic power leads to rising electricity costs. The reason for that, in essence, comes from the cost of ensuring that power is delivered whenever markets and people need it, and not when

The Cerebras "wafer-scale engine," measuring 8.5 inches on a side, is the world's largest chip, and is dedicated to computations prevalent in machine-learning forms of artificial intelligence. Photograph / Cerebras Systems.

nature permits it. It doesn't matter whether the reliability is achieved by maintaining what amounts to a duplicate, under-utilized existing grid (Germany's solution), or by using use more transmission lines and more storage. The results lead to far higher costs.

Policies that lead to higher costs and lower reliability for electricity will be increasingly in collision with the emerging demands for an AI-infused future. That may be the most interesting and challenging intersection of the worlds of bits and atoms. ◘

---

[i] https://simpleflying.com/henry-ford-aviation-pioneer-story/

[ii] https://semiengineering.com/ai-power-consumption-exploding/

[iii] https://www.technologyreview.com/2019/11/11/132004/the-computing-power-needed-to-train-ai-is-now-rising- seven-times-faster-than-ever-before/

[iv] https://www.bloomberg.com/news/articles/2023-03-09/how-much-energy-do-ai-and-chatgpt-use-no-one-knows-for-sure?sref=lHqvUqWg#xj4y7vzkg

[v] https://www.wsj.com/articles/artificial-intelligence-can-make-companies-greener-but-it-also-guzzles-energy- 7c7b678

[vi] https://arxiv.org/abs/1906.02243

[vii] https://www.scientificamerican.com/article/a-computer-scientist-breaks-down-generative-ais-hefty-carbon- footprint/

[viii] https://xcorr.net/2023/04/08/how-much-energy-does-chatgpt-use/

[ix] https://arxiv.org/pdf/1907.10597.pdf

[x] https://www.eetimes.com/qualcomm-targets-ai-inferencing-in-the-cloud/

[xi] https://fortune.com/2023/09/09/ai-chatgpt-usage-fuels-spike-in-microsoft-water-consumption/

[xii] https://www.wsj.com/articles/the-ai-boom-is-here-the-cloud-may-not-be-ready-1a51724d

[xiii] https://www.researchgate.net/publication/318461301_Green_and_Sustainable_Cellular_Base_Stations_An_Overview_and_Future_Research_Directions

[xiv] https://www.researchgate.net/publication/228774201_Power_Consumption_in_Telecommunication_Networks_Overview_and_Reduction_Strategies

[xv] https://www.us.jll.com/en/trends-and-insights/research/data-center- outlook?utm_source=Twitter&utm_medium=Social&utm_content=2631630870&utm_campaign=COMPANY+NE WSnon-specific

[xvi] https://www.nextplatform.com/2019/12/09/datacenters-are-hungry-for-servers-again/

[xvii] https://www.construction.com/construction-news/

[xviii] https://www.datacenterfrontier.com/cloud/article/11429232/the-eight-trends-that-will-shape-the-data-center- industry-in-2020

[xix] https://www.nytimes.com/2023/04/16/technology/google-search-engine-ai.html

[xx] https://www.scientificamerican.com/article/a-computer-scientist-breaks-down-generative-ais-hefty-carbon- footprint/

[xxi] Jevons, William Stanley, *The Coal Question,* Macmillan and Co., 1865. https://www.econlib.org/library/YPDBooks/Jevons/jvnCQ.html.

[xxii] Nordhaus, Ted. "The Energy Rebound Battle." Issues in Science and Technology, July 28, 2020. http://issues.org/33-4/the-energy-rebound-battle.

[xxiii] Larkin, Alice, Kevin Anderson, and Paul Peeters. "Air Transport, Climate Change and Tourism." Tourism and Hospitality Planning; Development 6, no. 1 (April 2009): 7–20. https://doi.org/10.1080/14790530902847012.

[xxiv] Roser, Max, and Hannah Ritchie. "Technological Progress." Our World in Data, May 11, 2013. https://ourworldindata.org/technological-progress.

[xxv] https://www.eetimes.com/iot-was-interesting-but-follow-the-money-to-ai-chips/?image_number=1

[xxvi] https://penntoday.upenn.edu/news/hidden-costs-ai-impending-energy-and-resource-strain

[xxvii] https://www.spglobal.com/commodityinsights/en/market-insights/blogs/metals/031723-southeast-asia-may- hold-the-key-to-chinese-aluminum-smelters-production- woes#:~:text=China%20is%20the%20world%27s%20largest,%2D%20and%20carbon%2Dintensive%20process.

NORTH DAKOTA'S POLYTECHNIC INSTITUTION

WHAT'S IN A NAME?

Douglas Jensen, EdD, President, Bismarck State College

In 2018, the North Dakota State Board of Higher Education designated Bismarck State College (BSC) as the state's first and only polytechnic institution. The North Dakota Legislature provided $38 million for the construction of a new polytechnic education center and for the start-up of new academic programs focused on meeting workforce needs. In December 2021, BSC announced the building of this facility, which is scheduled to open in late 2024 or early 2025.

The building of a polytechnic education center on the BSC campus was announced almost two years ago. The 88,000-square-foot facility will include state-of-the-art learning spaces such as a Security Operations Center (SOC), artificial intelligence and virtual reality labs, flex labs, a Digital Hive, collaboration spaces and a live-event venue. BSC broke ground in November 2022.

What is a Polytechnic Institution?

With only a small percentage of educational institutions in the country carrying this designation, many North Dakotans are asking, what is a polytechnic institution?

Simply put, a polytechnic institution is a university or college dedicated to the instruction of various technical arts and applied sciences using an educational model that engages students and industry with hands-on learning to develop workforce-ready knowledge, skills and degrees. This unique model allows students to combine multiple academic programs and skill sets to create customized college degrees embedded with industry-recognized credentials that are flexible and adaptable to meet the needs of both the student and industry.

A polytechnic institution combines the in-depth theory typically found at universities with practical career and technical education usually found at community or technical colleges. The focus is on offering a wide variety of certificate, associate and bachelor's degree programs with fewer admission requirements, smaller class sizes and lower tuition. Students advance their education and skills at their own pace and can even earn college credits and industry-specific certificates while in high school or the workforce. Those earned credits and certificates are stackable toward Associate of Applied Science and/or Bachelor of Applied Science degrees.

At Bismarck State College's 83rd Commencement Ceremony in May 2023, eight high school students graduated with a degree from BSC prior to graduating from high school, including Aurora Hill (above), a Bismarck High School junior, who walked across the stage to receive her Emergency Medical Technician degree.

Breaking Ground

In November 2022, BSC broke ground on the region's only polytechnic education facility. In fact, the closest peer institution is in Wisconsin. BSC's state-of-the-art center will feature project-based learning and nontraditional, hands-on, collaborative working environments.

"Having an institution align itself on the polytechnic mission, which is based on a hands-on, real-world, team-based learning, driven by the private sector, nothing can be more powerful," said North Dakota Governor Doug Burgum at the groundbreaking ceremony. Burgum has prioritized workforce development in the state. "If we have the workforce in North Dakota, the companies will come."

Designed by North Dakota-based ICON Architectural Group, the 88,000-square-foot polytechnic education center will connect to the west side of the BSC National Energy Center of Excellence (NECE) via a skywalk. The new facility will contain learning areas including a Security Operations Center (SOC); artificial intelligence (AI) and virtual reality labs; flex labs, where business and industry can partner on developing projects and ideas; spaces for operating and building new programming and for program equipment to advance applied research; a Digital Hive; collaboration spaces; and a live-event venue. Situated between the BSC Armory and NECE, this new facility will also have a spectacular view of the Missouri River.

SOC & Digital Hive

Inside the new facility, BSC will offer two unique digital technology environments with a mission to promote and advance the digital and gig economy: the Security Operations Center (SOC) and the Digital Hive.

The student-operated SOC is designed to train students in cybersecurity, giving them a near "real-world" hands-on experience as security analysts and security engineers under the leadership of a SOC manager. Students pursuing cybersecurity degrees will participate in the operation of SOC operations and debriefing activities after live operation events, maximizing their employability as cybersecurity professionals. Beginning with a basic design, the SOC will continue to mature, develop and expand based on the cybersecurity industry's needs—including future membership in the Multi-State Information Sharing and Analysis Center (MS-ISAC), which is the trusted resource for cyberthreat prevention, protection and recovery for U.S., state, local, tribal and territorial government entities.

The polytechnic education center will connect to the west side of the BSC National Energy Center of Excellence (NECE) via a skywalk. Situated between the BSC Armory and NECE, the facility features a beautiful view of the Missouri River. Cameras mounted atop NECE and the BSC Armory are capturing time-lapse videos of the construction progress, and the videos are updated monthly at bismarckstate.edu/FromTheGroundUp.

Also inside the new polytechnic education facility will be the Digital Hive, an environment designated to bring digital creators, designers, engineers and entrepreneurs of various backgrounds together to design and create future technologies. The Digital Hive's goal is to increase the entrepreneurship potential for the highly skilled, digital talent needed in our region. Open to the community as well as students, the Digital Hive will link talent to opportunity by creating network opportunities and providing a collaborative environment that embraces curiosity and fosters new ideas and innovation. The Digital Hive will empower students and community participants to explore interests, discover passions and strengthen their minds, thereby creating a talent pipeline to fuel the current and future economy, including the emerging gig economy.

Why Polytechnic Education?

The fundamental principle that sets a polytechnic institution such as BSC apart from traditional institutions is hands-on, applied learning. Here's a closer look at how learners benefit from a polytechnic education:

Flexible Learning Pathways

While traditional university courses focus on academic disciplines, such as English, math and history, polytechnic institutions focus on fusing technologies with applied learning in high-demand occupations in a variety of fields ranging from manufacturing, energy and agriculture to cybersecurity and health sciences. In addition to traditional undergraduate degrees, a polytechnic institution such as BSC also provides learners with opportunities to gain industry experience and earn industry-recognized credentials, certifications, certificates and degrees with apprenticeship options.

Polytechnic education is flexible. Many BSC programs share the same foundational core curriculum, giving students options to pursue short-term certificates, traditional associate degrees or customized stackable degrees. Part-time options, online/hybrid learning, and more evening and weekend programming provide even more flexibility for learners who want to upgrade their skills without sacrificing their personal lives or careers.

Flexibility is attractive to learners because each career pathway is customizable to meet each student's needs, offering a seamless transition from K-12 to higher education to industry at a considerable cost-saving.

Real-World Experience

It's not uncommon for colleges and universities to offer internship programs, but a polytechnic institution goes further by introducing students to real-world experiences early in their education—both inside and outside the classroom. BSC actively integrates work-based learning into all degree options and encompasses many opportunities in collaboration with business and industry, including service learning, project-based learning,

To ensure students are well-rounded and more marketable to employers, BSC focuses on giving students a solid foundation in STEAM, and the concepts from STEAM disciplines are woven throughout programs to align with business and industry demands.

The state-of-the-art polytechnic center will feature project-based learning and non-traditional, hands-on collaborative work environments and flex labs where business and industry can partner on developing projects and ideas.

apprenticeships, internships, field experiences and clinicals. Students apply their newly acquired skills through industry-sponsored projects that solve industry issues and improve efficiencies.

Because of this, learners at a polytechnic institution are workforce ready on graduation day. They have gained valuable experiences, acquired the career-readiness skills required to be marketable and graduate with a competitive advantage.

STEAM

STEAM education adds the arts to the traditional STEM foundation (Science, Technology, Engineering and Mathematics) as the integrated approach to learning at a polytechnic institution, which guides students' learning and opens their minds to the world around them. By incorporating the arts, STEAM embraces creativity, collaboration, critical thinking and communication, thereby nurturing curiosity and innovation.

To ensure students are well-rounded and more marketable to employers, BSC focuses on giving students a solid foundation in STEAM, and the concepts from STEAM disciplines are woven throughout programs to align with business and industry demands. One example of science intersecting with the arts within a program would be to offer a theatre course specific for health care students, which is under consideration. The course would provide students opportunities to role-play

The live-event venue is a versatile space that can be transformed for music and theatrical productions, class or community presentations, industry collaborations and much more.

different scenarios they might encounter with patients. All BSC programs—not just the technical programs—have a polytechnic influence, and faculty are incorporating STEAM education into their curriculum to help students develop highly employable skills in career-specific forms such as leadership, initiative, entrepreneurship, reliability, planning and organizing, critical thinking, problem-solving, communication and teamwork.

Industry 4.0

Today's students need to be trained in the ever-advancing digital technology that automates and improves business practices. The resulting transformations in business and industry—including the integration of increased automation, improved communication, and the production of smart machines that can analyze and diagnose issues without human intervention—is known as Industry 4.0, meaning the Fourth Industrial Revolution.

When computers were introduced into Industry 3.0, it was disruptive. Industry 3.0 was a giant leap ahead as the advent of computers and automation meant more robots were used to perform tasks previously performed by humans.

Today, the combination of cyber-physical systems, such as smart grids, autonomous automobiles, AI and the Internet of Things makes Industry 4.0 possible. The growing network of digitally interconnected smart machines creates and shares information, enabling business and industry to minimize waste and become more efficient and productive. This is the true power of Industry 4.0.

The Fourth Industrial Revolution differs from prior ones due to digitization's ongoing penetration into business and industrial operations, significantly increasing interconnectivity. Industry 4.0 standards are being integrated into all occupations in the workforce, and students must be equipped with the appropriate knowledge and skill sets. As a polytechnic institution, BSC ensures students meet these standards by embedding them into their curriculum and educational experience, whether in the classroom or experiential learning environments. Students graduate knowing how these automated processes work within their field and can adapt to the changes and improvements they bring.

Although Industry 4.0 will eliminate many jobs, especially highly repetitive tasks, it simultaneously creates many more new career pathways. The World Economic Forum's "The Future of Jobs Report 2020" predicted that by 2025, AI will replace 85 million current jobs worldwide—but create 97 million new jobs as all digital technologies require people with the right skills to conceive, build, program, maintain and repair.

Expert Faculty

Just as professors at colleges and universities, faculty members at polytechnic institutions, including BSC, have earned master's and doctoral degrees. The polytechnic advantage is that instructors also come to the classroom with many years of industry experience and credentials, meaning that BSC students gain valuable knowledge from career professionals in an applied-learning, project-based classroom environment.

Applied Research

Research conducted at universities tends to be based on theory and knowledge-sharing. Since polytechnic institutions engage and develop meaningful partnerships with industry leaders and stakeholders, BSC's research projects are designed to solve industry-specific, work-based problems, and the findings can be used to develop industry standards and processes and establish technical improvements that support industry growth and development.

Driven by Industry

As the state's only polytechnic institution, BSC is uniquely positioned to connect and engage with business and industry leaders to identify workforce needs and quickly respond with programs, certificates and even non-credit courses. At the core of BSC's polytechnic education model are 12 Business and Industry Leadership Teams (BILTs). The BILTs teams are composed of business and industry stakeholders who provide BSC with strategic advice and feedback on industry trends and the knowledge, skills and abilities required in future high-demand careers. This direct input influences the development of curriculum, program pathways and services.

"We've created BILTs made up of executive-level representation from local and state government, economic development entities, public education and regional business and industry to provide BSC advisory support," explained Alicia Uhde, BSC's Dean of Automation, Energy and Advanced Technologies.

As the state's only polytechnic institution, BSC is uniquely positioned to connect and engage with business and industry leaders to identify workforce needs and quickly respond with programs, certificates and even non-credit courses.

BSC has always responded to workforce needs with this industry-focused approach to education, which includes programs for specific skill sets and advanced training to benefit industry's current employees. Becoming a polytechnic institution has deepened and intensified this marriage of education and industry.

"There are more than 2,400 open jobs in Burleigh and Morton counties, and the positions are not all in one field or another. They are spread across blue-collar and white-collar industries alike," said Brian Ritter while serving as President and CEO of the Bismarck Mandan Chamber Economic Development Council. "BSC's evolution into a polytechnic institution enables us to address the community's workforce needs in a way that we've never been able to do in the past."

Because of their polytechnic-focused education, BSC students graduate workforce ready with the most up-to-date skills. They also benefit from the school's high employment rate due to BSC's focus on programs for high-priority, well-paying occupations.

Career Pathways in K-12

"I certainly believe that our partnership with Bismarck State College has greatly benefited Bismarck Public School students over the years, by allowing them to earn credits toward post-secondary certifications and degrees in career and technical education pathways while still in high school," said Dale Hoerauf, Bismarck Public Schools Director of Career and Technical Education. "This has been evident in the success of our Electronics articulation agreement and Med Terminology dual credit programs, to name a few. It is a win/win for students to be able to learn at their pace and receive post-secondary and industry-recognized microcredentials."

In partnership with K-12, BSC recently developed a Mobile App Development certificate and a Cybersecurity Fundamentals certificate, and both are available to students as early as sophomore year of high school. These are examples of certificates that can be "stacked" as college credits toward associate and bachelor's degrees.

"As the costs and debt associated with higher education continue to be prohibitive for many students and their families, offering college credit for career and technical education in high school provides a gateway to higher paying jobs, especially with companies actively looking to fill highly skilled positions," Uhde said. "These certifications are hands-on education that not only provide highly technical skills but also improve students' teamwork, problem-solving and communication skills."

Recently, BSC focused on expanding career pathways for high school students by increasing Dual Credit and Early Entry opportunities. The additional classes and certificates have resulted in a 46 percent increase in Dual Credit/Early Entry enrollments from fall 2021 to fall 2022 and a 108 percent increase from spring 2022 to spring 2023. And in May 2023, 41 students were classified as both full-time high school and college students, and eight graduated with a degree from BSC before receiving their high school diplomas.

"Dual credit fits perfectly within our unique polytechnic education model and offers students flexibility and opportunity. It is very possible for these 41 dual credit students taking a full college course load to graduate with their associate degree from BSC at the same time that they graduate from high school if they continue on this path," said Dan Leingang, BSC's Vice President for Academic Affairs. "This not only saves them time and money, it also puts them further ahead when they enter college and closer to meeting their career goal."

Career pathway expansion into K-12 will continue to be a focus of BSC's polytechnic education model.

Enrollment Growth

BSC recorded a 20.19 percent enrollment increase for the Spring 2023 semester, topping the fall enrollment, which had increased by 6.2 percent, for the same

academic year (2022-23) for the first time in more than 10 years. Then BSC saw an 8 percent increase for the Fall 2023 semester (4,065 students in total). The freshman class increased by 9 percent (2,433 students) to become the largest freshman class among the 11 institutions in the North Dakota University System (NDUS).

These statistics certainly go against enrollment trends in higher education, both in North Dakota and nationwide, and BSC believes this is the result of both an increased need for a skilled workforce and how the polytechnic model offers career pathways and flexible education options tailored to meet students' educational and career goals.

During the 2022-23 academic year, BSC also saw a 108 percent increase in dual credit enrollment. The five programs seeing the highest enrollment growth in Spring 2023 include Medical Laboratory Technician (45 percent), Electric Power Technology (41.75 percent), Agriculture Industry and Technology (34.29 percent), Cybersecurity & Computer Networks (22.13 percent) and Surgical Technology (24 percent).

Looking ahead, BSC anticipates enrollment will keep increasing as the polytechnic institution adds new academic programs in high-demand fields, such as agriculture, cybersecurity, energy, health sciences, manufacturing and automation.

What's in the Name?

As North Dakota's Polytechnic Institution, BSC prepares students to be workforce ready right out of college. Hands-on learning from industry trained faculty in unique education environments; the incorporation of STEAM and Industry 4.0 standards into curriculum, driven by industry experts; flexible learning pathways; internships with industry partners; and applied research opportunities, all together put students on a direct path to successful careers. ▣

Upper-level hallway at BSC's polytechnic education center with collaborative rooms on the left and the flex lab, shown on page 15, seen from above.

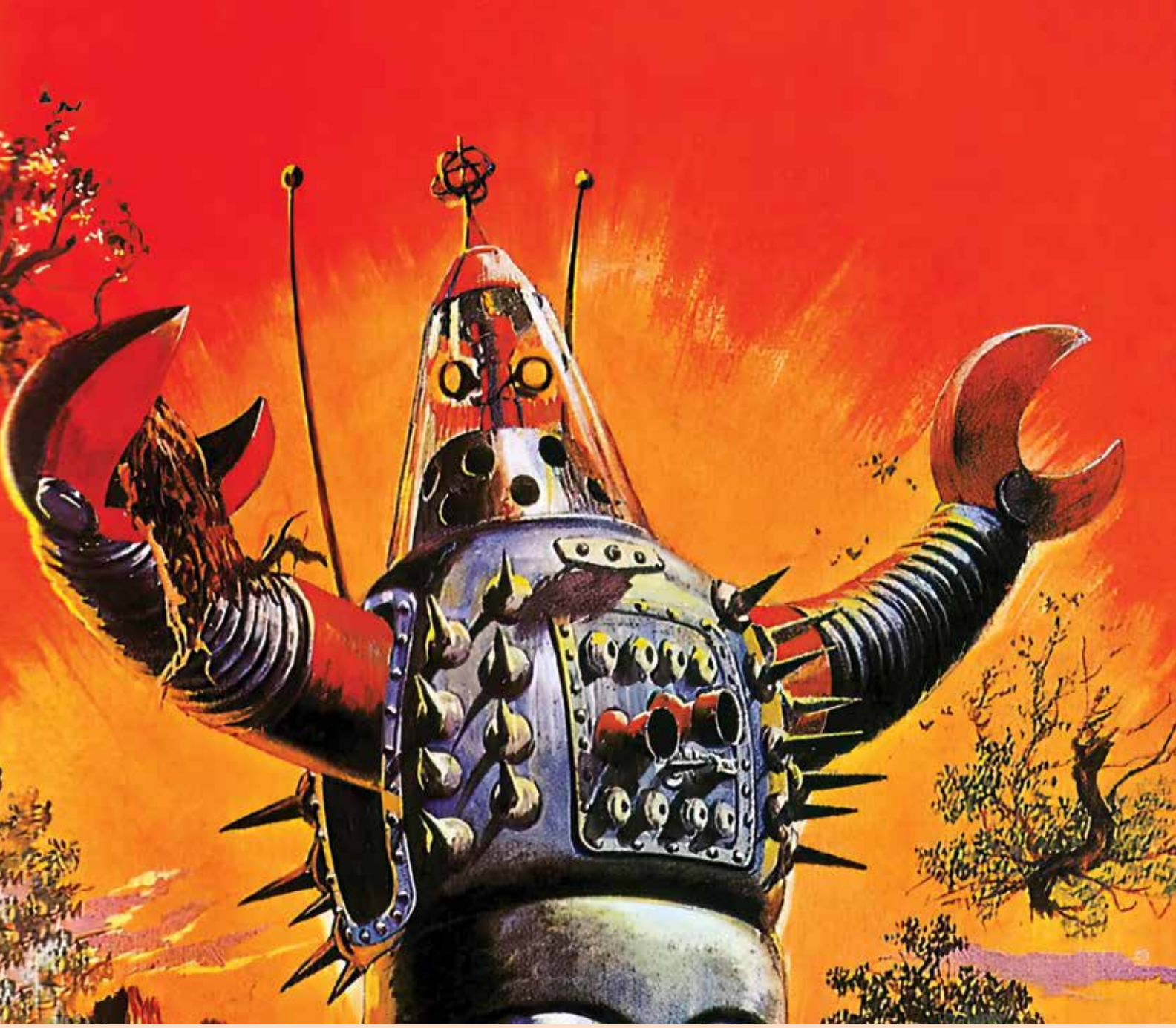


Illustration / Shigeru Komatsuzaki

OPINION

LOOKING BEYOND AI HYPE

To Maximize AI's Public Benefit

JEREMY STRAUB, PHD
Associate Professor,
Department of Computer Science, NDSU

Blizzards of hype surround artificial intelligence (AI) and threaten to prevent society from attaining its benefits. Concerns range from worries that AI might become Skynet from the "Terminator" movies—or an evil AI called The Entity[i] from the recent movie "Mission: Impossible-Dead Reckoning Part One"—bent on destroying humanity, to more down-to-earth concerns about job losses. Like any technology, some actors will use it for nefarious purposes[ii] and concerns about discrimination[iii] have also been raised.

In reality, though, AI is poised to bring massive benefits, including protecting us from cyberattacks, increasing our health,[iv] responding to emergencies[v] and even helping manage personal finances.[vi] In order to enjoy the advantages AI is poised to provide, we need to look beyond the hype—beyond calls for regulation coming loudly from a few computing luminaries—and focus on how innovation, technical discovery and entrepreneurship can be encouraged to enhance AI technologies and drive growth.

AI Isn't Out to Take Your Job

A big part of the hype surrounding AI is that it's going to cause[vii] large-scale job loss.[viii] Every technology that makes humans more efficient and can perform work that is currently done by humans changes the workplace. This isn't new. Concerns were raised about the cotton gin[ix]—a device that separated parts of the cotton plant, a burdensome task previously performed by humans—at the beginning of the industrial revolution. Printing presses, which required a time-consuming process of manually typesetting each page,[x] letter by letter—and numerous other technologies—were decried as a threat to jobs.

Of course, the reality is that, while some jobs changed and workers were displaced and moved into other jobs, there is, and was, no long-term mass unemployment caused by these (or other) technologies. In fact, technologies have increased Americans' standard of living by increasing the purchasing power of each hour of work. As Robert Tracinski aptly explains,[xi] the market forces created by new technologies' efficiencies increase the value of and demand for human labor.

Historical evidence, thus, provides a strong basis for an optimistic outlook, and history is the best guide we have to predict the future.

Protecting, Not Pernicious

The concern of AI being used as a weapon[xii] or by criminals is also commonly raised. This, though, is little different from concerns that might be raised about any tool. The same hammer that can be used to construct a building can also be used to smash a window. A backhoe can help create, or rapidly destroy, landscaping.

Because of its power, AI has, and will continue to be, used by governments, militaries, criminal organizations and numerous other entities. This, however, isn't a reason to try to stop AI through regulation. Quite the opposite. We need to avoid overregulation to allow those developing AI for positive and protective purposes to keep pace with those with criminal and other forms of villainous intent—as well as with nations seeking to use AI to assert dominance over us. Neither criminals nor foreign states are likely to be deterred significantly by our restrictive regulation of AI. In fact, they would benefit from it.

Regulated Already

The ongoing national discussion about AI regulation gives the impression that this technology is being developed and deployed in an unregulated "wild west." In reality, most concerns have already been addressed, which is why calls and efforts for regulation at the federal, state and municipal levels are more about political and commercial hype than substance.

Companies that might use AI for hiring are already covered by a variety of laws that prevent discrimination, for example. Regardless of whether an AI program or a human discriminates, the legal protections already exist. If anything, it will be easier to enforce laws against AI systems, based upon outcomes, as they make data-driven decisions and have no mechanism or incentive to try to cover up their conduct.

The same is true across numerous other areas of concern. Some laws may require limited modification —for example, to apportion liability for an AI's conduct between its developer and an operator.

However, these are minor changes to existing laws, not a new regulatory regime specifically created for AI.

Nor should we want AI to be regulated separately. Different standards for humans and AI activities, whether more or less restrictive, will inherently create loopholes that may prevent laws and regulations from being effective. For example, control of AI activities might be exempted from certain general regulations if AI-specific regulations exist, or general laws might be preempted by AI-specific laws. If these AI-specific laws can be subverted by small technical changes, which change how they are applied, or rendered obsolete by technological advancements, the regulatory intention of both sets of laws may not be achieved.

Well-written laws focusing on preventing harm and encouraging socially beneficial outcomes, for example, shouldn't need to be rewritten for AI. On the other hand, laws that focus intently on specific ways of preventing or producing outcomes might benefit from review and revision, even without considering AI.

Regulation Benefits the Big

Some of the companies developing AI are calling for regulation.[xiii] This is good for them and bad for everyone else.[xiv] Regulation transfers responsibility from companies to regulators and may remove or limit company liability for product-caused damage, if businesses can show they are following the regulations.

This is a recipe for irresponsible conduct. While some laws are needed to prevent developers from contractually avoiding liability for product failures, too much regulation can remove responsibility for AI failures, creating a moral hazard and promoting a lack of accountability.

AI regulation may also serve to limit the number of firms able to compete. New entrants, such as startups and other small businesses, may have difficulty understanding and complying with the regulations. They may also lack the financial and legal resources required to do so. Preventing new entrants into technology markets favors established firms while hampering technological advancement and disadvantaging society at large.

Regulation is Slow

Another important issue is the rapid pace of technological change. Regulations that focus on producing beneficial outcomes or preventing harmful ones might be helpful. However, those that take a more detailed approach to regulating technological design, development and operations will typically become outdated quickly but still remain in force. This could block innovation and undermine the development of societally beneficial technologies. Also, the longer regulations remain in force, the more likely it is that well-resourced firms will be able to invent workarounds to bypass the regulations, rendering them ineffective and subverting the lawmakers' and regulators' goals.

Liability Challenges

Instead of seeking to regulate AI separately, policymakers and lawmakers should focus on answering questions and clarifying laws about whether only humans can create works, make decisions and take actions. Questions abound, for example, regarding the protection,[xv] authorship[xvi] and ownership[xvii] of AI-generated intellectual property. These should be settled through a public lawmaking process that allows all concerned parties to be heard.

There is, similarly, a need to ensure a consistent and fair split of responsibility for AI's systemic failures. Software firms should not be allowed to transfer complete liability to users for the failure, acts and omissions of products that their users do not—and cannot possibly—understand fully, due to not having access to the underlying code and data.

However, in most cases, holding the AI developer solely responsible is not appropriate, since the configuration, implementation, prompting, lack of proper testing—and the decision to use AI at all for a given application—often rest in the hands of another party. These things can cause the system to fail, even without a defect. We also need to make sure that harmed consumers do not end up in the middle of a courtroom battle between AI developers and implementors that leaves the injured responsible for determining and proving which party is at fault.

Good Regulation

Regulation should focus on outcomes, such as promoting safety, preventing discrimination and protecting consumers. In each case, laws should require or proscribe outcomes and identify how these outcomes will be determined to have occurred. Specificity in conduct, though, is unhelpful. Imagine that we develop regulations outlawing murder that are narrowly tailored to certain weapons. We might proscribe murder by gun, knife or using a motor vehicle. If only these things are proscribed, a would-be murderer could bypass the law—and punishment—by simply choosing a different weapon, such as a baseball bat.

Laws may also be needed to ensure that records are retained to aid in the determination of responsibility. When humans commit acts or make mistakes, they (and other witnesses) can provide testimony. Since AI systems cannot be sworn in to testify, this will not be the case with AI decisions. It is, thus, reasonable to ensure that equivalent evidence, such as logs and other recordings of decisions and actions, is maintained. This will be especially important in determining fault among a technology developer, implementor and the possible contributory acts or negligence of an end-user.

We should also create regulations that help support technological development. One area of need is protection for the open-source community. We need to make sure that individual contributors to open-source projects cannot be held personally liable for contributions made in good faith that result in injury. Companies that benefit from free access to open-source projects—in particular, firms that repackage or use them to provide services to others—must assume the risk (and take action to mitigate it) of the free software they are utilizing. This protection is needed against both civil and criminal liabilities.

Ad Astra

AI has the potential to dramatically change our society for the better. It can help relieve humans of burdensome and repetitive tasks. It can improve and enable the personalization of entertainment options.[xviii] AI can help care for the sick and elderly.[xix] AI can aid the productivity and creativity of authors, directors and artists[xx]—and expand the possibilities available to them.

In order to deliver these benefits, AI must be allowed to grow in use and thrive. Regulations that prevent its deployment in order to protect special interests that lobbied for protection from technological advancement, regulations that place AI users at a disadvantage to those who use humans for similar tasks, and regulations that favor entrenched software developers are all contrary to the long-term public good. The next five years will be critical to human development and progress in numerous ways. One of the key decisions that each jurisdiction must make will be about how they treat AI.

These decisions may truly affect the proverbial 'fate of nations' with AI-embracers thriving and advancing, while AI-luddites find themselves left behind. ▣

---

[i] https://www.washingtonpost.com/technology/2023/07/28/mission-impossible-ai-not-realistic/

[ii] https://www.foxnews.com/world/hong-kong-arrests-6-loan-fraud-scheme-using-ai-deep-fakes

[iii] https://www.npr.org/2023/01/31/1152652093/ai-artificial-intelligence-bot-hiring-eeoc-discrimination

[iv] https://www.foxnews.com/opinion/i-love-ai-because-add-decades-our-lives

[v] https://www.foxnews.com/tech/ai-is-launching-911-call-centers-into-future-video-calls-triaging-redundant-reports

[vi] https://www.foxbusiness.com/technology/consumers-want-ai-help-manage-their-personal-finances-study

[vii] https://www.foxnews.com/opinion/artificial-intelligence-may-change-labor-market-but-doesnt-need-cause-long-term-harm

[viii] https://www.forbes.com/sites/jackkelly/2023/03/31/goldman-sachs-predicts-300-million-jobs-will-be-lost-or-degraded-by-artificial-intelligence/?sh=5631d3e6782b

[ix] https://www.asme.org/topics-resources/content/how-the-cotton-gin-started-the-civil-war

[x] https://www.history.com/topics/inventions/printing-press

[xi] https://www.discoursemagazine.com/economics/2023/06/26/says-law-of-robots-or-why-ai-wont-steal-all-the-jobs/

[xii] https://theconversation.com/artificial-intelligence-is-the-weapon-of-the-next-cold-war-86086

[xiii] https://www.nytimes.com/2023/05/16/technology/openai-altman-artificial-intelligence-regulation.html

[xiv] https://theconversation.com/does-regulating-artificial-intelligence-save-humanity-or-just-stifle-innovation-85718

[xv] https://www.jdsupra.com/legalnews/can-inventions-created-using-artificial-8457151/

[xvi] https://www.jdsupra.com/legalnews/no-copyright-protection-for-works-6704892/

[xvii] https://www.engadget.com/us-copyright-office-opens-public-comments-on-ai-and-content-ownership-170225911.html

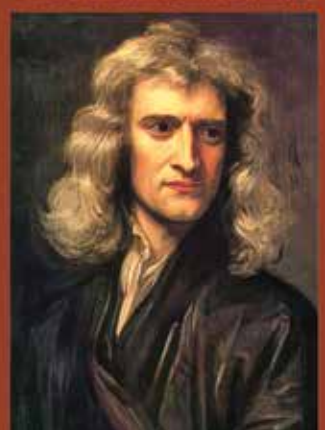
[xviii] https://www.usatoday.com/story/opinion/2023/05/10/wga-strike-pave-way-ai-generated-tv-movie-scripts/70198801007/

[xix] https://www.cnbc.com/2023/07/12/the-ai-revolution-in-health-care-is-coming.html

[xx] https://www.dailybreeze.com/2023/07/09/dont-crush-the-potential-of-ai-tech/

What Attorneys Should Know About Deepfakes

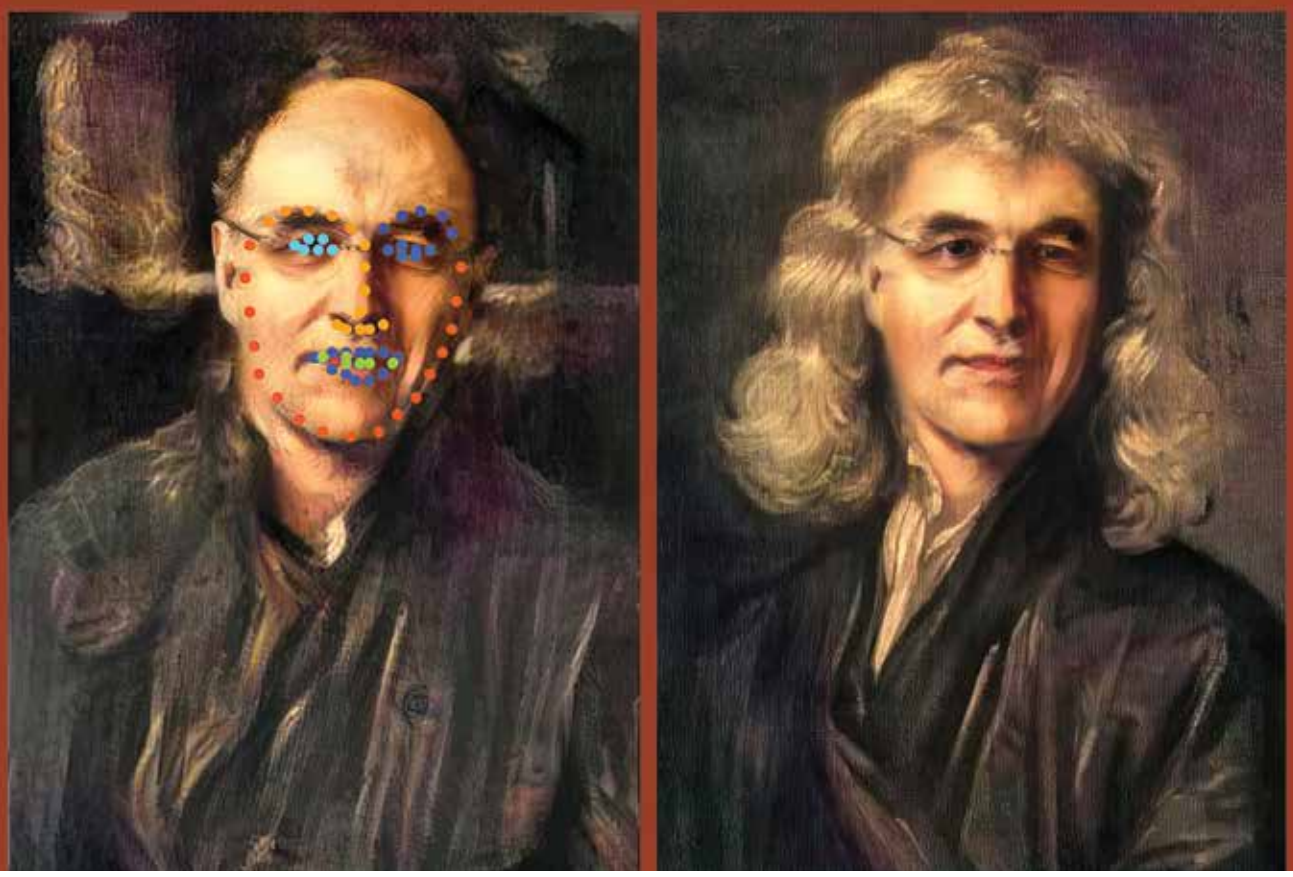
AI as Problem and Solution



BLAKE KLINKNER
Assistant Professor of Law
UND School of Law

Artificial intelligence (AI) is advancing at a startling pace, and society is grappling with AI's potential to be both beneficial and harmful. "Deepfakes" are one of the harms enabled by AI that has begun to show AI's potential to spread misinformation, sow distrust, and enable fraud and other criminal acts. Law and technology experts have also begun to sound the alarm on the threats, which deepfakes may pose to fair adjudications in courts of law, as AI has the potential to permit inauthentic evidence to be admitted at trial, while simultaneously allowing authentic evidence to be rejected based upon improper claims of inauthenticity.

In a deepfake, AI is used to create a new—and fake—image, video or audio, based upon a "sampling" of actual images, video or audio of a real person. For example, deepfake technology could scan the video of an actual political speech, delivered by a real politician, and then create a fake video purporting to contain a speech delivered by that same politician. The term "deepfake" is derived from the process used to create fake images, videos and audio, which uses "deep learning" algorithms that process real-life data (such as voice patterns and images of a real-life speaker) to then produce fake output (such as phony audio and video of that same speaker).



A face swap onto an original work of art using a neural net, which fit an internal model of one face and then apply it to the other. The parameters of the model are in effect learned from scanning lots of real-world scenes, and determining what's needed to reproduce them. Illustration / Stephen Wolfram

Deepfakes may be used for a variety of malicious purposes, with the common goal of tricking the public into believing that a person said or did something that the person did not actually say or do. Deepfakes have targeted politicians, such as one deepfake video purporting to show President Barack Obama launching into an obscenity laced tirade against President Donald Trump, and another deepfake distributed by a Belgian political party purporting to show a speech delivered by President Trump urging Belgium to withdraw from an international agreement.[i]

Deepfakes have also been distributed to vilify public figures and leaders of industry, such as a recent deepfake purporting to show Facebook CEO Mark Zuckerberg bragging about having "total control of billions of people's stolen data."[ii] Deepfakes have also been used to commit crimes. For example, one criminal scheme involved scammers, using deepfake technology to impersonate the voice of a relative, placing desperate calls to unwitting victims, pleading for the victims to quickly transfer funds due to a phony emergency.[iii]

Liar's Dividend

Legal experts predict that as deepfakes become more prevalent and difficult to detect, they will increasingly be the subject of evidentiary disputes in litigation. Deepfake technology has become easily accessible in recent years, and experts predict that parties will increasingly attempt to introduce evidence into court that is actually a deepfake.

Additionally, experts predict that legitimate audio, videos and images will increasingly be challenged in court as being deepfake in a phenomenon known as the "liar's dividend." According to the liar's dividend, as society "becomes more aware of how easy it is to fake audio and video, bad actors can weaponize" that skepticism. Because a "skeptical public will be primed

to doubt the authenticity of real audio and video evidence," actors can raise bad faith challenges by alleging that authentic evidence is actually deepfake.[iv] Consequently, if "accusations that evidence is deepfaked become more common, juries may come to expect even more proof that evidence is real," which could then require parties to expend additional resources to defend against unfounded claims that authentic evidence is fake.[v]

A recent high-profile example of a deepfake claim being raised in court to cast doubt upon an authentic video occurred in a wrongful death case pending against automaker Tesla, where the court rejected Tesla's assertion that a widely publicized video of CEO Elon Musk being interviewed at an industry conference in 2016 is a deepfake. The court found Tesla's assertion here to be "deeply troubling," and the court responded to Tesla's assertion by ordering a limited deposition of Musk on the issue of whether or not he made certain statements at the 2016 conference.[vi]

Detecting Deepfakes

Attorneys should be prepared to address deepfakes in their practices as deepfakes become more commonplace. The following are signs that a video, audio or image could be a deepfake:

Unreliable, questionable sources: Deepfakes are usually shared, at least initially, by unreliable, questionable, non-mainstream sources. If, for example, the originator of a videoed speech by a high-profile person is an unknown online entity, there is a strong likelihood that the recording is a deepfake.

Blurriness: In deepfakes, the target will often appear blurrier than the background. In particular, the hair and facial features of deepfake targets often appear blurry compared with other aspects of the video or image.

Mismatched audio: Deepfake visuals are often produced separately from deepfake audio, and then "stitched" together to create a final video. Consequently, visuals and audio can be misaligned, resulting in a mismatch between what is seen and what is heard. If, for example, there is a delay between what is heard and the movement of the speaker's mouth, such that it appears as though the speaker is lip-synching, this is a strong indication that the video has been deepfaked.

Mismatched lighting: Deepfakes will often retain the original lighting from the source video or image and transpose the original lighting into the new video or image, thus causing a mismatch of lighting within the final deepfake. If a video or image contains unusual, inexplicable shadowing, this is a telltale sign that it has been altered and might be a deepfake.

AI to Detect Deepfakes

As deepfake technology progresses, it will become difficult, and eventually impossible, for the human eye to detect deepfakes. Consequently, it will become necessary for attorneys to rely upon AI to detect deepfakes. Stated differently, we will need to rely upon AI to detect the works of other AIs, thus leading to an arms race between deepfake creators and deepfake detectors. In any event, attorneys should plan for a future in which they must safeguard against being fooled by deepfakes, be able to identify and counter deepfakes offered by their opponents in evidence, and be able to defend against bogus accusations that their own proffered evidence is deepfake.

This article was originally published in the June 2023 issue of Wyoming Lawyer.

---

[i] Ian Sample, "What are Deepfakes – and How Can You Spot Them?", The Guardian, January 13, 2020, https://www.theguardian.com/technology/2020/jan/13/what-are-deepfakes-and-how-can-you-spot-them; Hans Von Der Burchard, "Belgian Socialist Party Circulates 'Deep Fake' Donald Trump Video," Politico, May 21, 2018, https://www.politico.eu/article/spa-donald-trump-belgium-paris-climate-agreement-belgian-socialist-party-circulates-deep-fake-trump-video/

[ii] Von Der Burchard, *supra* note 1.

[iii] Pranshu Verma, "They Thought Loved Ones Were Calling for Help. It was an AI Scam," Washington, March 5, 2023, https://www.washingtonpost.com/technology/2023/03/05/ai-voice-scam/

[iv] Shannon Bond, "People Are Trying to Claim Real Videos are Deepfakes. The Courts are Not Amused." NPR, May 8, 2023, https://www.npr.org/2023/05/08/1174132413/people-are-trying-to-claim-real-videos-are-deepfakes-the-courts-are-not-amused

[v] *See id.*

[vi] "Elon Musk's Statements Could Be 'Deepfakes,' Tesla Defence Lawyers Tell Court," The Guardian, April 26, 2023, https://www.theguardian.com/technology/2023/apr/27/elon-musks-statements-could-be-deepfakes-tesla-defence-lawyers-tell-court

Grand Farm

is a network of growers, technologists, corporations, startups, educators, policymakers and investors working together to solve problems in agriculture with applied technology. In early 2022, Grand Farm was awarded a $10-million matching grant by the North Dakota Legislature and Department of Commerce to further advance agriculture technology through a world-class Innovation Campus.

For information, please visit:

GrandFarm.com

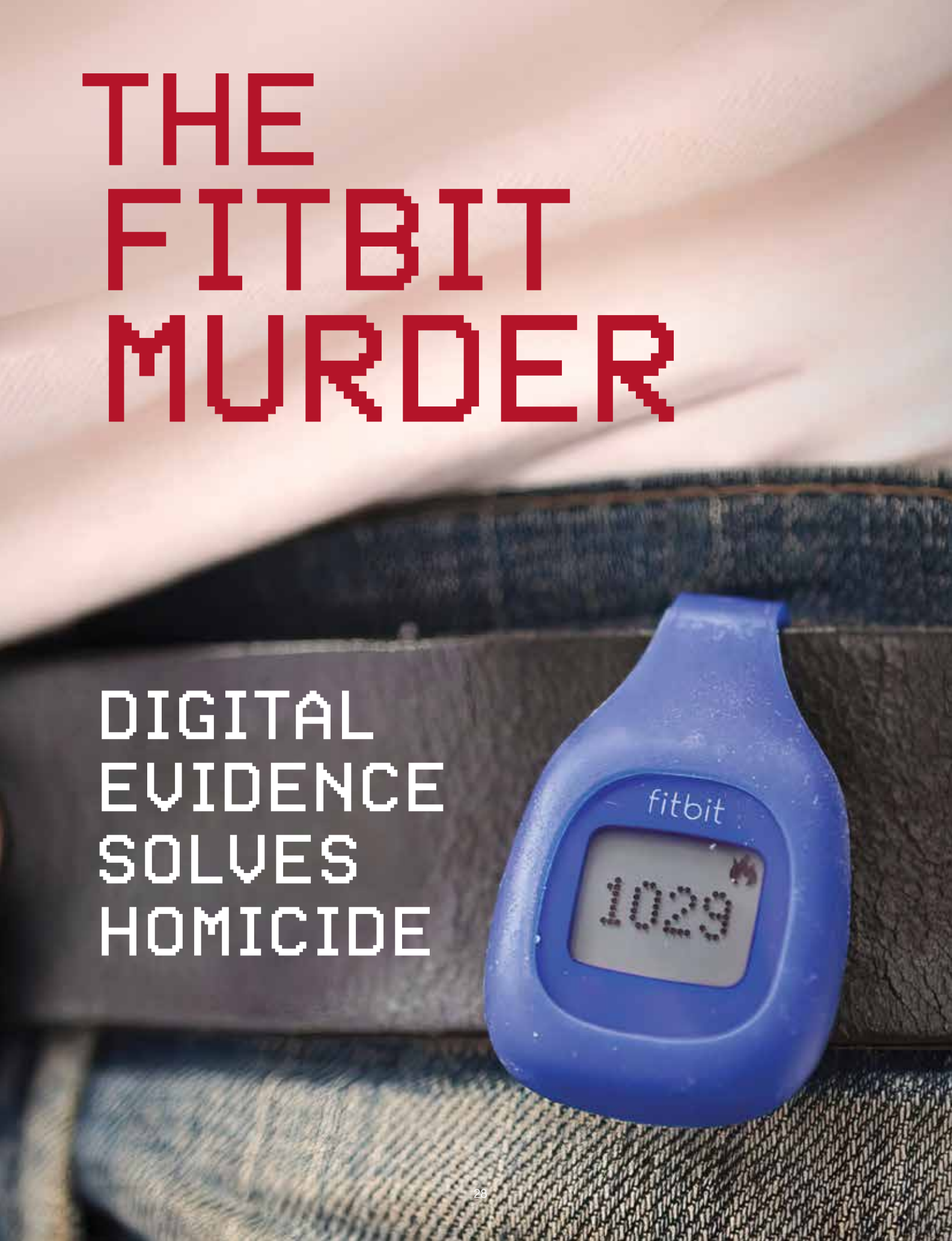
- **Grand Farm's Innovation Campus** will bring together researchers, growers, industry, startups and government agencies to ideate about and execute innovations aimed at solving some of the world's largest challenges in agriculture.

  The campus will provide expanded acreage for the deployment of agriculture technology projects, rapid prototyping capabilities, and increased research and educational capacity.

- **The Grand Farm Innovation Shop,** rendered above, will be the first building to go up on the Innovation Campus. The shop will be utilized, as examples, for workshops, education, equipment storage, and as event and project spaces.

THE FITBIT MURDER

DIGITAL EVIDENCE SOLVES HOMICIDE



ARICA KULM, PHD

Director of Digital Forensics Services
Dakota State University

Two days before Christmas in 2015, Richard Dabate called 911 to report that his home had been broken into and his wife shot. When the police arrived, they found Dabate on the floor of the home's main level, partially bound to a folding chair with zip ties. His wife was in the far corner of the basement lying dead, with bullet wounds to the stomach and head. The story that would unfold is one of a marriage that was very different on the inside than how it appeared on the outside.[i]

Richard and Connie Dabate were married on July 4, 2003. The couple settled in Ellington, Connecticut, and had two sons, Richard ("RJ") and Connor, who were 9 and 6 years of age in 2015. Richard, aged 40, was a computer technician and Connie, aged 39, worked as a pharmaceutical sales representative for Reckitt Benckiser and was the family's main "breadwinner." The Dabate family lived at 7 Birchview Drive, a four-bedroom, colonial-style home at the end of a long driveway in an affluent neighborhood.

Friends describe a happy couple living a seemingly idyllic life with the normal marital spats involving disagreements over money. The digital timeline discovered by investigators leading up to and following Connie's murder would tell a much different story.

On the morning of December 23, 2015, the Dabates' sons went to school, and Richard said he left for work between 8:30 and 8:40 a.m., only to realize around five minutes into his trip that he'd forgotten his laptop at home. After pulling over to send a quick email from his phone to his boss to let him know he'd be late, he turned around and headed back home. Upon arrival, which according to Richard was between 8:45 and 9:00 a.m., he heard a noise upstairs and found a man in camouflage and a mask, with a voice like Vin Diesel, rummaging through his wife's jewelry in a closet. The intruder made demands including money and credit cards with pin numbers, which Dabate said he handed over, even though the intruder didn't bring a weapon with him.

When Richard heard the garage door and then the kitchen door open, he assumed his wife had returned home early from her exercise class at the Indian Valley YMCA in Ellington. He yelled for her to run, but rather than escaping the house, she ran to the basement. The intruder, Richard claimed, followed Connie there, struggled with her and then shot her in the back of the head and in the stomach with a gun Richard had purchased two months before, but it had never been fired. The gun, a Ruger .357 Magnum revolver, was found to have no fingerprints—not from the intruder, Connie or even Richard. How the intruder supposedly got possession of the gun was never clarified as Richard's explanation changed several times.

Richard was then stabbed in the legs with a box cutter, which the intruder also found in the house, and then he tied Richard to a chair with zip ties. The intruder began burning Richard with a blowtorch, but when a struggle ensued, Richard managed to burn the intruder's face before he ran off.

During the initial on-site investigation, police used dogs to try to pick up the scent of the intruder, but they failed to pick up an exit trail of any intruder. The first dog kept circling back to Richard and once tried to jump into the back of the ambulance where Richard was being treated for his injuries. Two additional police dogs failed to pick up any scent of an intruder leaving the property.[ii]

Connie Dabate

Many parts of Richard's story didn't make sense from the beginning. There were no signs of forced entry into the home, nothing was taken, and none of the neighbors saw anyone suspicious in the neighborhood. Richard claimed to have struggled with the intruder during the invasion but the closet where the struggle was supposed to have taken place was found to be neat and tidy, drawers closed and undisturbed. The intruder was allegedly rummaging through jewelry when Richard encountered him, however the jewelry was still in the closed drawers in the closet. Richard also had no bruises on him despite claiming he had wrestled with the intruder and claiming that the intruder had used "pressure points" to subdue him and tie him to the chair.[iii] All weapons used by the intruder during the attack originated inside the residence. There were also no signs of a struggle on Connie's body.

The Other Woman

Investigators would uncover a story worthy of a soap opera from Richard about another woman. She was a high school friend with whom he had begun having an affair in 2014 after she divorced her husband. Richard claimed that he and Connie wanted another child but were unable, so his friend was going to serve as a surrogate, and then the three of them would raise the baby together. After discussing artificial insemination, they decided to go the more traditional route and achieved a pregnancy. Later, he would reveal that his wife did not know about the affair or the pregnancy, which he then said was unplanned.

Digital Evidence

Investigators used multiple items of digital evidence, detailed below, to piece together a timeline showing that Richard was lying about many things. Put together, the digital evidence told a tale of deception, marital strife and ultimately murder.

> Cellphone Evidence

Richard gave written consent to allow the search of his iPhone. Connie's iPhone was located in the right-side waistband of her sweatpants, under her jacket facing the floor, and was seized at the crime scene.

Recovering data from a cellphone is not always easy during investigations: The make and model of phone, version of installed operating system, condition of the device, and whether it is locked with a passcode or PIN code, all determine whether data can be recovered. Even with a search warrant, suspects are not compelled to reveal their passcodes, and in the case of a deceased victim, his or her passcode might not be known to those left behind.

In this case, however, investigators were able to recover the data from Richard's and Connie's cellphones, revealing several incriminating items of evidence.

A year prior to her murder, in December 2014, Connie made two entries in her iPhone Notes application, entitled "Why I Want a Divorce," detailing the reasons she wanted to divorce her husband. These reasons included the irresponsible way Richard handled money, such as taking funds from accounts that didn't belong to him, for being an unfit parent, for his uncaring attitude toward her, for not coming home on time and that he "acts like a kid constantly." She also had a list in the same Notes app of good things about him, which was much shorter. Connie's cellphone also revealed a text argument between her and Richard the day prior to the murder, in which she accused him of lying about a cable bill.

Richard's iPhone also revealed important information, including a text to his girlfriend two days prior to the murder promising to divorce his wife. As well, there were several alarm notification text messages during the morning hours indicating the arming and disarming of the home alarm system.

> Camera Evidence

The data from video surveillance cameras at the YMCA where Connie went for exercise classes was analyzed during the investigation. Video surveillance equipment can be of many different makes and models and store data in different ways. Some have internal hard drives in a main unit that store the videos captured by the attached cameras, others have individual memory cards within each camera, and some send the data to cloud storage. It's important on any system to verify the date and time setting on the camera. Often video surveillance cameras are set up and forgotten. Then power outages or daylight savings time changes, as examples, can cause the date and time to be different from the actual date and time. In this case, the date and time on the YMCA video surveillance cameras

The former home of Richard and Connie Dabate at 7 Birchview Drive, Ellington, CT.

were accurate. Cameras from the parking lot showed Connie arriving at about 8:53 a.m. She soon found out that her class was canceled and was seen on the cameras leaving at approximately 9:08 a.m.

Social Media Data

Social media data is often analyzed during investigations to build a pattern of behavior and relationships. Connie's Facebook records, as well as those of Richard's girlfriend, were analyzed.

Connie sent a private message from her cellphone via Facebook Messenger at 8:58 a.m. to a psychotherapist requesting an appointment to be hypnotized "because there's a lot going on right now." (Altimari, 2017)

From 9:40 to 9:46 a.m., Connie posted two videos on Facebook using her iPhone and then posted a message to a friend, again through Facebook Messenger. The Internet Protocol (IP) address used to access Connie's Facebook account was assigned to the couple's house. Nothing in Connie's Facebook data, either publicly posted or in her private messages, indicates a divorce pending, or that she had knowledge of Richard's extramarital affair or the pregnancy.

The analysis of Richard's girlfriend's Facebook records further told the story of their relationship, as she confided in friends via Facebook Messenger about her love for him and his promises to divorce his wife. Information from a divorce attorney interviewed during the investigation revealed that Richard had met with him on June 30, 2015, but the attorney was not retained.

Computer Evidence

Modern computers can also present challenges to investigators, similar to cellphones. Proper forensic technique, when examining a hard drive, requires investigators to create a forensic copy of the hard drive, verify that it matches the original with hashing,[iv] and then perform the analysis from the copy, leaving the original evidence unchanged. Many computers now come with encryption enabled by default. If a drive is encrypted, a copy of that drive can still be made. However, if it is encrypted and the computer's password is unknown, all you have is a copy of a hard drive that is also encrypted, and the data cannot be read without the proper decryption code or key. In addition, many modern computers and tablets do not have a removable hard drive, making it less straightforward when obtaining a forensic copy.

In this case, Richard's Microsoft Surface Pro tablet was able to be examined and revealed items of interest in his web-search history, including visits to websites during the time he alleged he was driving to work on the morning of December 23. At 8:26 a.m. and again at 8:27 a.m., there were visits to Facebook.com. At 8:37 a.m., there was a Google search for "long-lasting tattoo ideas." At 8:41 a.m., a login to Richard's Outlook email account was made from the tablet using the IP address from within the home. There were also several visits to other sites, including the Indian Valley YMCA at 9:18 a.m. to download a "Group Exercise Schedule." Two minutes later, he searched the ESPN website for the "Mike and Mike" show, which was the last time he used his computer that morning.

Microsoft Corporation Records

Account access, such as those from Microsoft or other online providers, can aid in tying a login event and sometimes a device to a location. Connections to the internet, such as that in the Dabate home, are assigned an IP address, which is a number that uniquely identifies that connection to the Internet Service Provider. Service providers such as Microsoft, which provide products including email, Cloud storage and conferencing apps, have logs of when users connect to their products and from which IP addresses. Richard's Microsoft account records were analyzed and showed that he accessed his account several times that morning from his home IP address. In this case, the records did not indicate which specific device accessed the records, but the logins located in the Microsoft logs correspond with activity found on either his cellphone or laptop computer. Records indicate an email was sent from Richard's Outlook email address to his supervisor at 9:04 a.m. Richard said he sent this email while in his car after pulling over almost two miles from the home. The typical reach for a home internet router is only 300 feet outdoors with no obstructions. Richard could not have sent the email from his car as claimed.[v]

Fitbit Murder: Timeline of Events

July 4, 2003	Richard and Connie Dabate were married.
December 2014	Connie made entries on iPhone Notes app on reasons to divorce.
June 30, 2015	Richard met with a divorce attorney.

DECEMBER 23, 2015: THE MURDER OF CONNIE DABATE

8:26 & 8:27 AM	Visits to Facebook.com from Richard's computer.
8:30 to 8:40 AM	Richard claimed he left for work.
8:37 AM	Google search for "long-lasting tattoo ideas" from Richard's computer.
8:41 AM	Login to Outlook.com from Richard's computer (and several additional visits between 8:41 & 9:20 AM).
8:45 to 9:00 AM	Richard claimed he returned home to encounter and struggle with an intruder. Then Connie returned home and was shot by supposedly by the intruder after a brief struggle.
8:46 AM	Connie's Fitbit indicated she likely left for the YMCA.
8:47 AM	Richard logged into the home alarm website attempting to disarm the alarm.
8:50 AM	Richard successfully disarmed the alarm system, then armed it again from the home alarm website.
8:59 AM	Richard disarmed the home alarm system from his keychain fob.
8:53 AM	Connie arrived at the YMCA (camera evidence).
8:58 AM	Connie sent Facebook private message to her psychotherapist.
9:04 AM	An email is sent from Richard's Outlook email account to his supervisor at work. He later claimed he sent it when he pulled over in his car, but the email was sent from his home IP address.
9:08 AM	Connie departed from the YMCA (camera evidence and Fitbit inactivity consistent with driving).
9:18 AM	Visit from Richard's computer to the Indian Valley YMCA website.
9:20 AM	Visit from Richard's computer to ESPN's "Mike and Mike" show website.
9:23 AM	Connie's Fitbit registered activity when the home's garage door opened (alarm data).
9:40 to 9:46 AM	Connie posted two videos to Facebook.
10:05 AM	Connie's last movements registered on Fitbit.
10:12 AM	The panic alarm for the home security system was activated from Richard's keychain fob.

Fitbit evidence

Connie was wearing a Fitbit One activity tracking device, which is designed to be worn clipped to a waistband (as it was on Connie) or other article of clothing, rather than directly contacting the skin. According to the Fitbit website, the data tracked by this model includes steps taken, floors climbed, recent activity levels, distance traveled and calories burned. The data resets at midnight, depending on the time zone selected, with the data being displayed in a user dashboard when it is synced.

Connie's Fitbit indicated she likely left for her YMCA fitness class around 8:46 a.m. There was a period of inactivity consistent with her driving there, and then activity consistent with her actions while at the YMCA. The next period of inactivity began at 9:08 a.m., the time she was recorded leaving the YMCA. The next recorded movement on Connie's Fitbit was at 9:23 a.m., the same time as the alarm system registered that a door between the garage and the kitchen was opened. The Fitbit records showed that her last movements inside the home were at 10:05 a.m. —nearly an hour after Richard told detectives she had been killed by the masked intruder.

Richard's story was that Connie had come into the house and ran directly to the basement after encountering the intruder. However, during the time Richard alleged this occurred, the Fitbit recorded Connie walking a distance of 1,217 feet. This was much further than the approximately 125 feet between her car and the basement, conflicting Richard's account of Connie's return home.[vi]

Home Alarm evidence

At 10:12 a.m., the panic alarm for the security system was activated from Richard's keychain fob. It was the only time the panic alarm went off that morning. Richard had claimed when he turned around to go home to retrieve his laptop, he received alerts about the alarm and emailed his boss that he was returning home to check on it. Richard's cellphone text messages showed that these various notifications, which included the arming and disarming of the system, were received that morning.

Richard Dabate at trial sentencing.

The data recovered from the home alarm system indicated it was armed and disarmed several times that morning. Experts testified about how the system was functioning properly and about the various features of the alarm system, including two keychain fobs that had a maximum range of 500 feet, which would be shortened if any objects were between the fob and the alarm system, or if cellular interference were to intervene. The key fobs were listed as "Keychain" and "KEYFOB6." KEYFOB6 was on the same keyring as Richard's Nissan vehicle key fob.

The system also contained motion sensors designed to detect body heat. The alarm company's records show movement around the house that is inconsistent with Richard's account of the events. The alarm system was armed as "stay," indicating people can be in the house and not trigger the alarm, at 8:47 a.m., when Connie was at the gym and Richard claimed to be on his way to work. The alarm system was armed as stay from the keychain fob on Richard's keychain, which would have required someone to be at the house. At that same time, 8:47 a.m., Richard also logged into the website of the alarm system from his phone to attempt to disarm the system, which was unsuccessful.

Later, at 8:50 a.m., Richard was successful in disarming the system from the website using his phone. Then also at 8:50 a.m., less than a minute later, he armed the system from the website. At 8:59 a.m., the system was disarmed from Richard's keychain fob. All of this took place during a time that

Richard stated he wasn't at home but rather on his way to check the alarm notifications and to retrieve his forgotten laptop. In addition, the first door to open and close after the system was disarmed at 8:59 a.m. wasn't an outside entrance, as one would think for a person returning home to retrieve a lost item as he claimed. Rather, the basement door opened and closed indicating whoever opened and closed the door had to already be inside the house. At 9:23 a.m., the opening of the garage door coincided with the data recovered from Connie's Fitbit activity at the same time, indicating she arrived home from the YMCA.

Also noted was that Richard cancelled his subscription with the alarm company 12 days after his wife's murder, despite claiming that the home broken into resulted in Connie's demise.

> Verizon Cellphone Evidence

Records returned from cellular carriers typically include data, such as subscriber information, call records and text message records. Verizon cellphone records for Richard's girlfriend were seized, and they showed that the text messages between her and Richard were exchanged with Richard using a Google Voice number rather than his primary cellphone number—a fact he failed to disclose during his interview with investigators. Those messages were not stored within the messages associated with Richard's primary cellphone number so were not recovered from the search of Richard's cell phone. Had the girlfriend's cellphone records not been obtained and only the messages from Richard's cellphone analyzed, the texts between Richard and his girlfriend may not have been located. This highlights the importance of gathering multiple sources of digital evidence.

Richard's Verizon cellphone records were also obtained. These showed text messages and the notifications received from the alarm company.

Trial & Conviction

The murder of Connie Dabate became known as the "Fitbit Murder," due to the use of the Fitbit data in helping to solve the crime. However, there were multiple items of digital and other evidence pieced together by investigators to aid in Richard's arrest and ultimate conviction for his wife's murder. The 22-day trial began on April 5, 2022, in Rockville Superior Court, and included 600 exhibits and 130 witnesses. The trial was interrupted by the COVID-19 pandemic after the first jury selection was almost complete in March 2020. Then Richard's attorney died in June 2021. The jury selection resumed in February 2022 with a new jury selection due to the amount of time that had passed, with some of those selected having moved away in the intervening months, in some cases out of state.[vii]

On May 10, 2022, Richard was found guilty of the murder of his wife, tampering with evidence and making false statements. Prosecutors in the case asked for a 60-year sentence, which was surpassed by the judge after hearing from Connie's family and friends. It took nearly seven years for her murder to be resolved and justice to be served, but in the end, Richard Dabate was sentenced to 65 years. Currently, he resides in the MacDougall-Walker Correctional Institution in Suffield, Connecticut.[viii]

---

[i] Note that much of the information in this article is taken from (Arrest Warrant Application, 2017).

[ii] Zymaris, E. (2922). "Troopers Detail How K9s Led to Richard Dabate as Murder Suspect." Retreived from News 8 wtnh.com: https://www.wtnh.com/news/connecticut/tolland/troopers-detail-how-k9s-led-to-richard-dabate-as-murder-suspect/

[iii] Altimari, D. (2017, 04 23). "A Marriage Marked by Secrets, A Murder Case Months in the Making." Retrieved from Hartford Courant: https://www.courant.com/2017/04/23/a-marriage-marked-by-secrets-a-murder-case-months-in-the-making/

[iv] Hashing is explained in Arica Kulm's previous article in Dakota Digital Review, "Solving Crime Through Digital Evidence," available at: https://dda.ndus.edu/ddreview/solving-crime-through-digital-evidence/

[v] Mitchell B. (2020, November 5). "What Is the Range of a Typical Wi-Fi Network?" from Lifewire: https://www.lifewire.com/range-of-typical-wifi-network-816564

[vi] H-11 Digital Forensics. (2017, 09 14). "Fitbit Used as Key Evidence in Murder Case." Retrieved from H-11 Digital Forensics: https://h11dfs.com/fitbit-data-used-as-evidence-in-murder-case/

[vii] Leavenworth, Jesse. (2022, 02, 28) "Jury Selection in Connecticut 'Fitbit' Murder Trial to Begin—Again." Retrieved from Hartford Courant: https://www.courant.com/2022/02/28/jury-selection-in-connecticut-fitbit-murder-trial-to-begin-again/

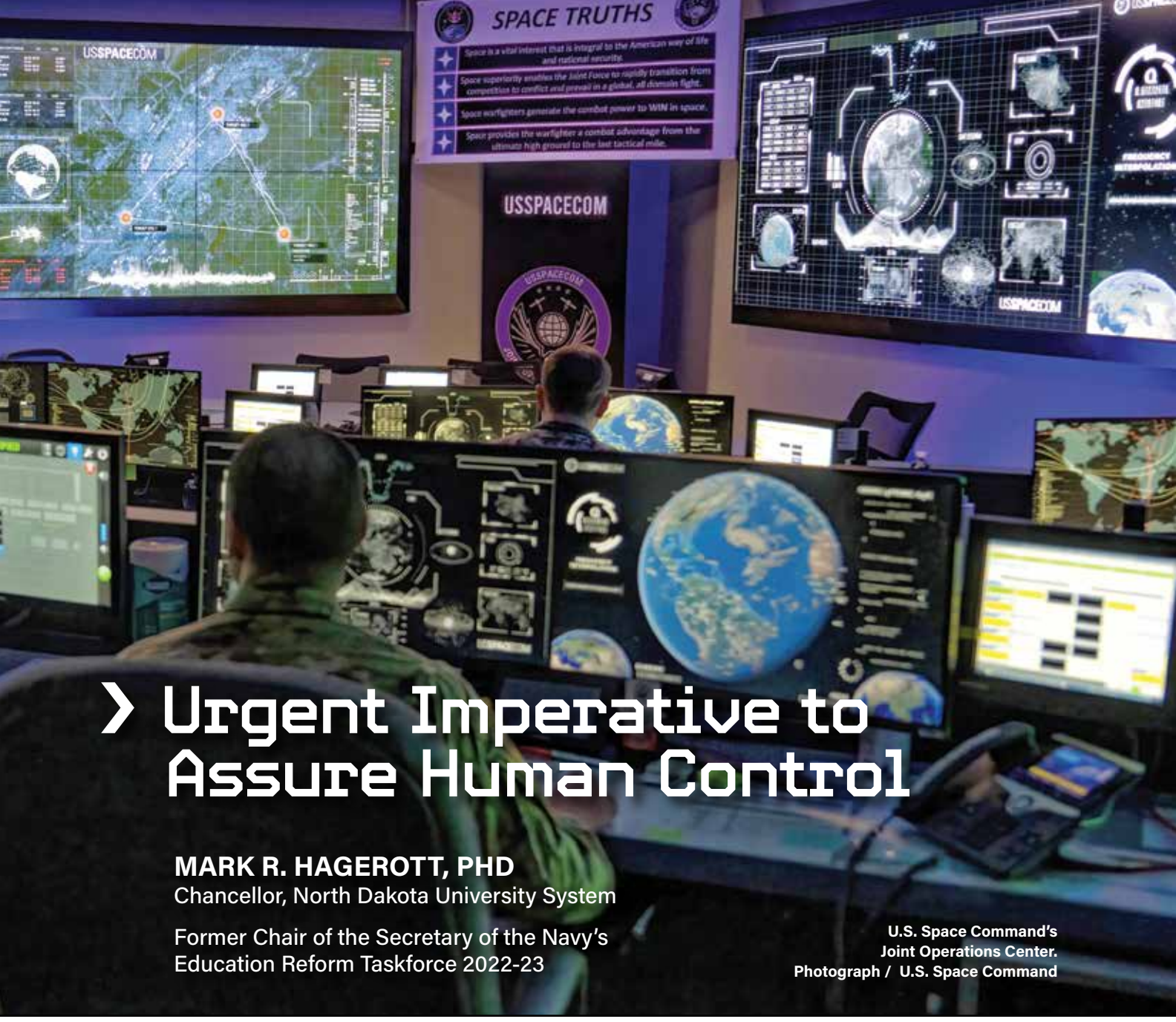
[viii] Connecticut State Division of Criminal Justice. (2022, 08 18). Retrieved from Connecticut's Official State Website: https://portal.ct.gov/DCJ/Press-Room/Press-Releases/08182022DabateSentencing

Limits to Automated & AI-Controlled Military Systems

The history of technological improvement is not a straight line, and the most thoughtful political leaders have always questioned the promises and underlying assumptions of their engineers and scientists.

As artificial intelligence (AI) surges to new levels of capability, this generation of political leaders should be asking our technical elites perhaps the most important questions ever: Have we compromised trust and resilience for speed and efficiency in our rush to digitize? As the technological elites automated and accelerated military and business processes, uploading more and more data into the Cloud, what vulnerabilities have been created now and for future generations if these decisions gain inertia and the moment to reconsider has passed? Is it possible that by the gradual accretion of thousands of uncoordinated decisions to automate and digitize, we collectively have increased the possibility of massive, systemwide hacking of our digital systems, thus creating brittleness and the possibility of an existential threat to American security?

The views expressed in this article are the author's alone and do not represent those of the U.S. Navy, the Department of Defense or the State of North Dakota.



Urgent Imperative to Assure Human Control

MARK R. HAGEROTT, PHD
Chancellor, North Dakota University System

Former Chair of the Secretary of the Navy's Education Reform Taskforce 2022-23

U.S. Space Command's Joint Operations Center. Photograph / U.S. Space Command

This article proceeds on the supposition that a decades-long unquestioned orthodoxy surrounding digitization and automation in pursuit of speed and efficiency has undermined national resilience and created potential existential security vulnerabilities. Such loss of resilience is not fated but a choice emanating from an identifiable military philosophy of armed conflict that originated in the U.S. military and later spread to government and industry. Understanding that the roots of our challenge are philosophical, and that *humans have a choice* to restore resilience and trust, allows us to identify solutions, however radical they may seem, in the existing speed-efficiency paradigm. That solution includes shaping, narrowing and maybe limiting where we allow digitization and AI to come together, but also includes a radical reemphasis on human agency, human education, training, skills and abilities, to include the substantial elevation of human control in existing human-machine teams. Stuart Russell, a world-leading AI expert, when pondering the coexistence of humans and AI, concluded that a reemphasis on the human factor was critical, and that only human cultural change, akin to "ancient Sparta's military ethos," could preserve human control and agency in the Age of AI.[i]

› War-Game Epiphany

Headlines worldwide now proclaim news of accelerating AI capabilities. Task forces are being convened from the White House to the Pentagon to Wall Street to better understand the threats and opportunities this presents.[ii] Perhaps AI's most important aspect, potentially existential to national security and critical state and local systems, is the convergence of AI and cybercrime and cyberconflict. But there is a fundamental problem of epic proportions that is being ignored, perhaps purposely, because the solution set might be so radical.

The problem is that, quite simply, there is *no technical solution set* that alone can assure human control of AI and its increasingly integrated technologies when these are used to power cybercrime and cyberconflict. The radical solution set must include a profound rethinking of *human* knowledge, skills and abilities, as well as the preservation of tools to assure ultimate human control of digital systems in the face of hacked AI algorithms. The solution set may have to include shaping, narrowing or limiting the reach of AI and digitization. How did I come to this radical insight? And why hasn't such a solution, costly and neo-Luddite as it might seem, been adopted already?

A decade ago, I saw evidence that a techno-philosophy was gaining unquestioned adherents in domains where hacked AI vulnerability could emerge with portentous and unpredictable consequences. I was attending a war-game exercise sponsored by the Office of Defense, Research and Engineering in the Office of the Secretary of Defense (OSD) that examined the future of technology, conflict and war.[iii] About the same time, I was asked to present at the Geneva Convention on Certain Conventional Weapons conference on lethal autonomous weapons systems in the spring of 2014.[iv] Based on the war games and discourse at Geneva, attended by more than 100 ambassadors, it became clear that three transformative technologies would challenge the Department of Defense (DOD) to its core: AI-powered autonomous killing machines; a next generation of ever-more ubiquitous digital, AI-influenced communication networks; and, as anticipated in science fiction, enormous pressure would build for increased human-digital machine integration.

But I also took away something else from the games: The sense that the *momentum*[v] of the DOD's R&D/acquisition system was propelling the U.S. military toward a strategic conundrum of historic proportions, that the pursuit of speed and efficiency would create brittleness and undermine trust and resilience that could extend well into the future and become almost irreversible. The three major transformative technologies, mentioned above, progressively replace human decision-making, knowledge, skill and physical abilities with intelligent digital devices (IDDs) that are vulnerable to cyberattack.[vi] In a pre-cyber conflict age dominated by the U.S., such a replacement of the human with machine might produce all positives: reduced risk to Americans, faster flow of information, higher-performing battlefield units, more efficient state infrastructure and perhaps cost savings. But in the face of rising cyber powers, this proliferation of IDDs in communication networks, robotics and human-decision aides and enhancements may place our national defense and state/local security at risk.[vii]

Rapid digitization and AI emergence in confluence with cyberwar argues for what may seem counterintuitive but has historical precedent. During this period of uncertainty, DOD, state governments and key infrastructure corporations should slow deployment of digital-AI programs that displace human skill and decision-making and should slow the retirement of mature, stand-alone technologies. In parallel, DOD and the states should reestablish selective training programs to preserve or regain critical human-centric knowledge, skills and abilities.

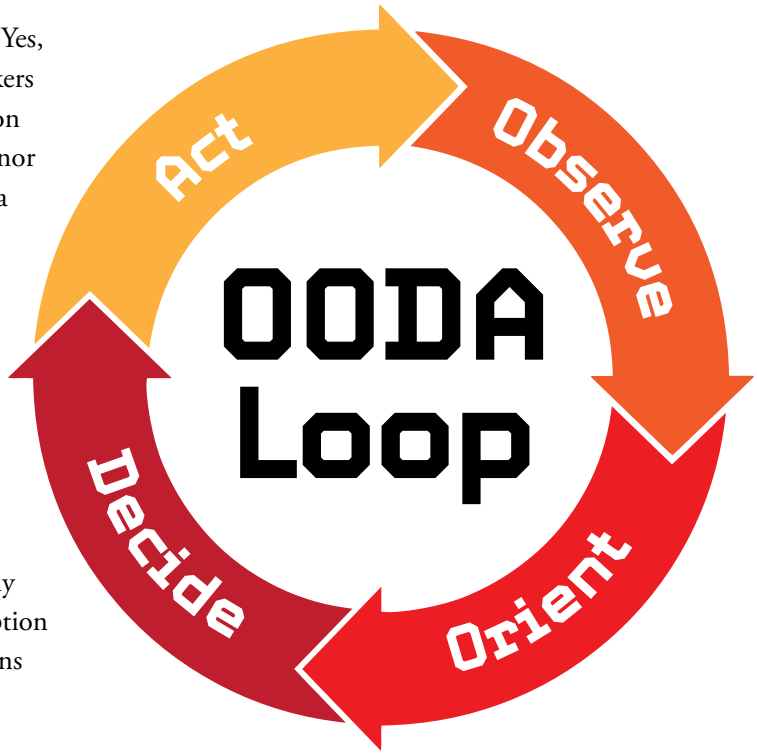
> **[T]here is *no technical solution* set that alone can assure human control of AI and its increasingly integrated technologies when these are used to power cybercrime and cyberconflict.**

But can the U.S. military change this trajectory? Yes, but first both political and military decision-makers must understand that the trajectory of automation and AI application is a human choice, not fated nor inevitable. Yet we are up against a growing inertia of blind acceptance. How did we get here? And what are the origins of this particular philosophy?

› Over-Automation & Privileging Speed

So, what shaped human thinking that we have privileged digital speed and efficiency at the potential cost of trust and resilience? A philosophy so doctrinaire that it was embedded as an assumption in doctrine and even war games where assumptions should have been tested. The futuristic four war games I attended examined the evolution toward unmanned systems, ever larger information networks and electronic human-machine integration. It was often argued that such electronic-based systems could get inside an enemy's decision cycle and give us an advantage in what is known as the OODA (Observe, Orient, Decide, Act) Loop. And, I was persuaded: Unmanned systems with AI processors could compute faster in many cases than a human; computer-enabled tactical electronic communications systems could transmit more data faster than the human voice or non-computerized communications; and soldiers aided by yet ever-more electronic and web-enabled devices could allow fewer, lesser trained humans to do more tasks faster than personnel without these devices. Many of these are already in the field, including handheld GPS linked to iPads reducing reliance on human navigation skills on land or sea, and computerized translation programs that, while convenient now, will ultimately reduce the incentive for soldiers to maintain natural human-language proficiency.

But where human action and decision-making are displaced by IDDs, new questions of security arise, now known as cybersecurity. Its close relative, cyberpower, enables an actor to use computer code to take control of, influence or degrade another actor's IDDs or communications systems.[viii] Our



The OODA (Observe, Orient, Decide, Act) Loop was developed in the 1990s by USAF Colonel John Boyd as a military strategy focusing on agility (by facilitating rapid, effective reactions to high-stakes situations) to overcome an opponent's raw power. The OODA Loop has also been applied successfully to business and industry, and more recently shown applicable to cybersecurity and cyberwarfare.

country continues to proliferate hackable IDDs in an increasing number of systems based on the implicit assumption that the U.S. will maintain information dominance and thus a favorable cyber balance of cyberpower. The assumption underlines DOD's race to build fleets of unmanned vehicles, build ever-more complex and netted electronic information systems, and deploy ever-more electronic decision aids to our sailors and soldiers.

But is it reasonable to assume that our increasingly automated and computerized systems are and will remain cybersecure, trustworthy and resilient?[ix] I think that several of these suppositions are or will very soon be in doubt for a simple reason: Unlike more traditional forms of physical power, cyberpower relationships can shift unpredictably and leave our nation in a condition of relative uncertainty. Thus our ability to predict, observe and react may be inadequate to maintain information dominance and cyber superiority. Why is this so?

› Uncertainty of Cyberpower

Cyberpower calculations are increasingly opaque, and as a result, determining which country is or will remain in the cyber lead is uncertain.[x] Unlike security calculations and arms races of the past, where counting tanks, battleships or ICBMs provided a rough measure of relative technological power, such calculations are more difficult if not impossible today. The addition of each new IDD to the already millions of such devices in the DOD inventory adds another conduit for cyberattack and contributes to rising complexity.[xi] Due to the proliferation of IDDs, we are on a trajectory towards the time when nearly all critical systems and weapons may be accessible and hackable by computer code. In this new electron web of machines, if one of our stronger cyber rivals gains a strategic computing advantage (perhaps a breakthrough in supercomputing or cryptography), the consequences could range from the tactical to the strategic across our netted systems and automated platforms to the detriment of soldiers who have become dependent on electronic devices.

And, there is a dawning revelation of the vulnerability of automated and remotely piloted vehicles. DOD's Defense Advanced Research Projects Agency (DARPA) some years ago instituted the High-Assurance Cyber Military Systems program (HACMs) to provide better protections to the American drone fleet. Most recently, DARPA all but admitted it was struggling to keep up with the pace of AI evolution and began a series of workshops, AI Forward, in the summer of 2023, to bridge the fundamental gap between the AI industry and DOD.[xii] The U.S. Air Force Research Laboratory Chief Microelectronics Technology officer admitted that, "At a high level … our program offices and our contractors do not have good visibility into the electronics and designs that they're actually delivering into the field … . If you don't know what is in your system, how can I possibly trust it?"[xiii]

It is not just the scope that's concerning but also the speed at which power can shift. Espionage and treachery have been historical realities since before the Trojan Horse. But with the growing reliance on IDDs, automation and networks, the costs of failure are accelerating, magnified and broadcast systemwide. Edward Snowden released documents cataloguing National Security Agency (NSA) activities and did significant damage to U.S. national security, but the operating military forces were largely unaffected. What if critical electronic and automated systems were either hacked or compromised? With ever-more netted, automated military Supervisory Control and Data Acquisition (SCADA) systems, the ability of a hacking to disable ever greater segments of our infrastructure might be possible and happen rapidly—perhaps with little warning.[xiv]

In the recent past, it took several years to build battleships or nuclear submarines to change a naval balance of power, during which time we could see the shift in power coming, for it was hard for a potential adversary to hide the 50,000-ton behemoths. Conversely, today, with our growing reliance on automated machines, an opponent might gain a strategic cyber advantage with the changing loyalties of a single programmer or the breakthrough by a team of programmers producing powerful algorithms, all occurring with a minor physical footprint, perhaps in a non-descript office building, all in a relatively short period of time.

› Root Cause: OODA & Trust/Resilience

How did the frontline military become so dependent on automation and electronics? Short answer: In the 1970s and 80s, leading military thinkers and technologists privileged speed as a determinant on the tactical battlefield. Electronics and automation accelerated Colonel John Boyd's OODA Loop, which made sense in the environment in which they developed their ideas. Given how the problem was defined then, automation with its ever-higher speeds of decision-making made good tactical sense, especially when pilots were independent and not networked.[xv] But a combination of group think among military leaders and growing momentum in the R&D/acquisition system now propels us toward greater automation, lesser direct human control, even though the environment and conditions have changed.[xvi]

As mentioned previously, Boyd's concept of "getting inside an enemy's OODA Loop" gained a growing

number of adherents across all the military services, in all platform communities, and eventually in the government and industry writ large.[xvii] Boyd was a fighter pilot who built on his experience in the cockpit and argued that the central objective of new systems was to speed up the decision cycle: to "Observe, Orient, Decide and ACT (shoot)" an enemy first. However, in the OODA Loop, the trustworthiness of the decider (the pilot) was not a factor. The number of American fighter pilots who became traitorous in the cockpit has been minimal, possibly none, after a century of air combat experience.

To be sure, there have been American and Allied combatants who turned traitorous on the battlefield. But in cases where enemies infiltrated or swayed a decision-maker, typically in the ground forces, the effects were localized. Such human hacking cannot, by the nature of human-analog functioning, ever become systemwide. Witness the isolated "green on blue" attacks (in which an Afghan soldier(s) attacks American troops) that periodically occurred in Afghanistan, which could not spread systemwide at the digital speed of light. In contrast, where there are IDDs run by ones and zeroes, the entire netted system becomes vulnerable, whether squadrons of Unmanned Aerial Vehicles (UAVs) or a division of soldiers and their electronic decision aids. A cyberattack could thus turn a system of UAVs against us or corrupt a platoon's kit of electronic decision aids and produce strategic or tactical losses on multiple battlefields simultaneously.

Thus, the "more speed" orthodoxy has been turned on its head with the possibility of cyberattack and questions of trust and brittleness. With continued proliferation of IDDs—without keeping cyber superiority—the inevitable and unavoidable result will be increased surface area for cyber vulnerabilities, higher chance of cyber penetration and a reduced resilience. Is this orthodoxy on the wane? Recently, DOD adopted another OODA-Loop driven initiative, the Joint All Domain Command and Control (JADC2) to include integration with nuclear command and control systems.[xviii] Though this strategy document tweaks the phrasing ever so slightly to "Sense, Make Sense and ACT," make no mistake, this is OODA-Loop philosophy arguing for faster systems integrated even with nuclear control systems.

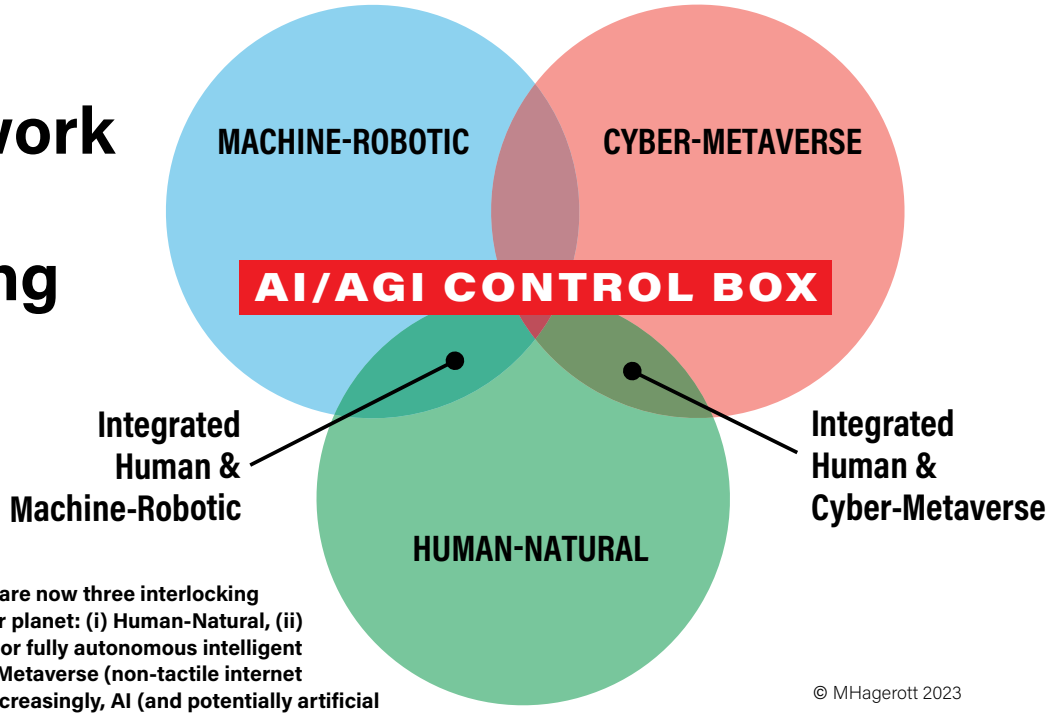
To persuade a critical mass of military officers to question their preference for speed and automation requires an alternate philosophy, that I suggest is trust and resilience. Trust is well understood, but what are key features of resilience for combat units?

Resilience, the ability to sustain a cyberattack and then restore normal operations,[xix] is an especially important consideration in America's cyber cadres. But maintaining resilience with frontline units, which navigate in dangerous waters or face a kinetic military environment, is different than maintaining resilience in relatively static civil systems and different than systems within safe borders, such as Fort Meade and NSA. In systems situated safely behind secure borders and walls, a momentary cyber breach, termed "zero day," can be sustained typically without physical damage (STUXNET-like attack not withstanding). The coding problem will be corrected, and if data was lost, there typically exist backup locations for critical data.

But when a system is physically located in a potentially kinetic or environmentally dangerous military area, a zero day relating to navigational, electrical, propulsion and/or defensive systems may be unrecoverable. A cyberattack in conjunction with kinetic attack can result in irreparable harm that cannot be regained with a software patch or by accessing the backup data. Thus, viewed from this perspective, our growing dependence on millions of military IDDs, on growing numbers of unmanned and more highly automated systems in operating forces, in conjunction with the rise of both AI and cyber warfare, has changed the

A cyberattack could thus turn a system of UAVs against us or corrupt a platoon's kit of electronic decision aids and produce strategic or tactical losses on multiple battlefields simultaneously.

Visual Framework of Our Digitizing World



With digitization, there are now three interlocking realms of activity on our planet: (i) Human-Natural, (ii) Machine-Robotic (near or fully autonomous intelligent robots) and (iii) Cyber-Metaverse (non-tactile internet including the Cloud). Increasingly, AI (and potentially artificial general intelligence [AGI]) is taking control of realms (ii) and (iii) and the convergence of all three, as well as threatening to usurp most of the Human-Natural realm. This graphic illustrates the challenge and urgency of limiting the reach of AI to reduce cyberattacks and long-term human deskilling.

problem and requires new thinking. As well, there are potential major consequences of inaction, not just for the current generation but the next. As mentioned previously, technological systems have a well-documented tendency to 'lock in' the decisions and choices of the first generations of users. If we do not thoughtfully engage this question of AI, digitization, automation and cyber resilience now, the vulnerabilities may become endemic for our children.

› What should be done?

Ten years have lapsed since my first uneasy feelings at the war games. With the explosion in AI and advancement of cyber tools, the stakes are even higher now, as General Mark A. Milley, Chairman of the Joint Chiefs of Staff, opined in his farewell address: "The most strategically significant and fundamental change in the character of war is happening now, while the future is clouded in mist and uncertainty."[xx] But Milley is not alone. The National Science Foundation has also sounded the alarm. The clouds of digital data and millions of robotic or automated machines may reach new levels of efficiency if guided by AI, but such a combination also poses a risk that "these [AI] systems can be brittle in the face of surprising situations, susceptible to manipulation or anti-machine strategies, and produce outputs that do not align with human expectations or truth or human values."[xxi]

If deliberate policy does not slow the trajectory of R&D/acquisition, the culturally privileged OODA philosophy will drive even wider proliferation of AI and IDDs in new systems,[xxii] resulting in further automation and human-deskilling of combatants and a loss in resiliency.

There are several courses of action to defend against rising cyber uncertainty and the unique challenges of military and civilian resilience. Of course, we will never disinvent the digital computer nor eliminate our many networks, and we shouldn't try. In cases where speed is critical, we must risk deploying highly automated systems, even with the attendant cyber risk. DOD is already pursuing the most technically sophisticated courses of action: investing in the latest cyber-secure systems,[xxiii] contracting for the best software provided by the best cybersecurity firms; R&D to create near unhackable IDDs or near unbreakable cryptography (if a breakthrough

in quantum computing doesn't first break current encryption); secure, government-controlled production facilities; development of unhackable inertial navigation systems; hiring more and better qualified computer network defense experts;[xxiv] incorporate the most rigorous antitamper technologies in the growing fleet of unmanned systems. But developing computer code to check on computer code is a costly, and some say an impossible task, a problem identified by Alan Turing and now known as the Halting Problem—when does the computer know it can halt searching for a virus? For these reasons we must carefully shape, narrow or limit the application of more powerful and vulnerable AI to existing and future digital systems.

But there is another way to increase cybersecurity, to reduce the surface area of cyberattack, increase resilience, detect failures and provide unhackable code that will carry out orders uncorrupted: Keep more humans in the loop, making decisions and using their natural human skills to control machines and communicate when the automated systems show signs of failure or corruption.

What does rebuilding human-machine resilience entail? We should maintain more legacy and humancentric systems, as well as modify and reform officer and enlisted education to ensure our operators can navigate, fight, command and control (C2) in a less-netted and computer-aided environment. This includes training to restore human agency in the key functions of move-shoot-communicate, to include both maritime and land navigation; non-computer aided communication; non-networked, human-directed warfighting capabilities; human language proficiency and critical thinking—before these KSAs (Knowledge, Skills and Abilities) have been totally supplanted by AI-robotics or dangerously atrophied. Rebalancing the roles of AI-powered digital machines and ensuring the vitality of the human member of the human-machine team might cost more but could preserve human agency and create a more resilient human-machine system, reducing the surface area for cyberattacks and the possibility of cyber-silent failures, as well as provide the ultimate unhackable code: natural human cognitive and physical processes.

The following are military areas of concern, where the OODA-Loop fueled rush to replace-the-human with faster, increasingly autonomous systems has opened the door for digital and AI-related cyberattack.

› Education to Rebuild Human Skills

As stated before, we cannot unplug the DOD from all AI and digital machines. Rather, the goal is to consciously decide how to shape AI and automation to ensure a more trustworthy resilient human-machine team in cyber conflict or war. Preserving or invigorating the human element in the human-machine team will bring advantages, including reduced surface area of possible cyber penetration; increased resilience; and less likely silent failures, since more trained and skilled operators will be available to perceive when equipment performance begins to degrade, a gap which was demonstrated in the STUXNET cyberattack.[xxv] To this end, combat training should include conditions of degraded communications, in which control of forces would have to be conducted solely with human-readable and human-audible transmissions, similar to the EMCON strategies during the Cold War.[xxvi]

Why does this convey a possible advantage? Humans remain all but impervious to cyberattack. To disable a human and his/her human-operated mechanical system typically requires the physical destruction of such a system, person or platform. Thus, the continued manual human presence in key military processes may in fact increase the resilience of our systems.[xxvii]

To its credit, the Army officer training programs continue to require proficiency at land navigation, unaided by GPS or any digital device. While costly, the sea services may need to consider a reinvestment in the older radio navigation systems, which were all but disestablished in the past decade, and human-centric celestial navigation should be preserved.[xxviii] Easy access to GPS position data now results in deskilling human operators,[xxix] who grow yet more reliant on electronic systems in a reinforcing cycle. A recently released report concerning the grounding and destruction of the USS Guardian explicitly notes the

crew's overreliance on GPS data and digital charts, and their failure to use physical/optical verification (the *eyes* of the deck officers) to avoid shoal water.[xxx] To their credit, the U.S. Naval Academy in Annapolis reversed a 15-year-old decision to eliminate celestial navigation, and now the entire student-body receives basic instruction and training.

DOD should continue to require new officer candidates to be electronic-enhancement free, and DOD should continue to avoid the implantation of any electronic devices in military personnel to ensure that natural senses remain acute.[xxxi] On the battlefield, DOD should proceed carefully in providing soldiers or sailors with external electronic sensory or decision aids, which may in the short term provide a memory boost, facial recognition, language translation capability—all desirable tools in a cyber-secure environment. But such technologies will inevitably lead to the deskilling of humans and increasing cyber vulnerability while reducing a unit's resilience. The DOD is aware of the growing inventory of 'wearable sensor tech' and should be applauded for their recent efforts to study this issue.[xxxii]

And not just the military educational institutions have a responsibility, but also the entire national K-12 system. The threat of society-wide disinformation and overreliance on the internet and now AI is growing. A key counterstrategy to preserve trust and resilience in our young military ranks is a reemphasis on critical thinking both in K-12 and collegiate programs, including ROTC and the service academies. To that end, there must be an expanded national effort, from kindergarten to the doctoral level (Cyber PK-20), to impart a basic understanding of digitization, cyber hacking and AI, tantamount to basic math and language literacy efforts of the 19th century. Similarly, federal and state governments should consider a new Digital Service Academy and even a national Digital-Cyber Land Grant Act to invigorate digital and cyber education at the collegiate level.[xxxiii]

› Shaping Deployment of AI & Automation

Our culture—at DOD and throughout America—privileges the new and technical. But we misinterpret our history if we think that previous successful technological revolutions proceeded without abatement or delay. Many technical revolutions proceeded in fits and starts, as new technology was tested, found wanting and then reapplied with greater success. In the past, many experts were convinced, as examples, that the neutron bomb and nuclear-powered aircraft were wonderful ideas. But these and other technological applications proved unwise, although submarine and large-ship nuclear propulsion and civilian power generation were widely adopted to great benefit. Going further back, speed and labor saving were not the unquestioned policy drivers as today.

In the 19th century, the purchase and employment of speedy steam ships was delayed in favor of more resilient, reliable, steam-sail hybrids.[xxxiv] In the nuclear-power revolution of the mid-20th century, Admiral Hyman G. Rickover purposely chose a relatively costly, highly trained human-centric organizational approach over a labor-saving, more computer dependent system of reactor operations, display and control—a choice that the Chernobyl nuclear meltdown incident confirmed as profoundly wise.

A strategy of slowing or narrowing the deployment of AI and automation has now gained national support.

Poseidon

Russia's Poseidon (also known as Status-6 and, in NATO, as Kanyon) is a large, intercontinental, drone torpedo, which is nuclear-powered and nuclear-armed. Poseidon can travel autonomously undersea up to 6,200 miles to attack enemy coastal cities. Illustration / Covert Shores

Most recently, hundreds of scientists, engineers and entrepreneurs called for a six-month delay in the further fielding of advanced AI algorithms, until the larger implications of such technology could be studied.[xxxv] But the issue goes deeper and further than just the latest general AI algorithms. Some potential use cases are described below.

Large Ship Navigation & Russian Drone Nuke Sub

The U.S. Navy is under strict human control and exhibits great resilience in contrast to the now desperate Russian attempt at dangerous and reckless automation. For several decades, debate has swirled around the possible Soviet-Russia development of a 'dead hand,' or highly automated nuclear retaliatory Doomsday Machine. This was never confirmed until recently: The Russians have developed a nuclear-armed drone, which is an unmanned submarine capable of cruising several thousand miles at high speed with the mission to destroy coastal cities. The euphemistic characterization of this as a "torpedo" stretches any accepted use of the term, as torpedoes were always of limited range and tactical. This is an example of loss of human agency and profound loss in resilience: Putting a nuclear weapon on a submerged drone using AI and satellite navigation, both of which abandon human control and open the path for AI poisoning, hacking or even self-hallucination.[xxxvi]

Might the U.S. Navy eventually over-automate in a rush to keep pace with other rising powers? In the high-speed missile battlefield, OODA-Loop speed will most likely remain a necessity. Years ago, the Navy committed to high-end automation as a solution to missile attacks and built the AEGIS self-defense system, which allowed a robotic, lower-level AI to take control of a ship's weapons. But humans were present, as I can attest as a former combat systems officer on an AEGIS ship. I observed the low-level AI computer's independent action, and I was able to turn the analog key to disengage the firing signal, thereby shutting down the robotic system.

Yet there are theorists who argue the benefits to depopulating entire ships, and indeed the U.S. Navy is on track to develop a fleet of unmanned ships and submarines. But as both AI- and cyber-hacking tools grow in sophistication, does the cost-benefit calculation of an increasingly robotic, AI-powered fleet begin to change? If we fight against a nation that gains even temporary cyber superiority, our ships may be at increased risk of navigational data corruption, they may be compelled to slow their speed of movement while they await the outcome of the cyber battle. Again, this could pose a profound risk.

AI or Human or Both in the Cockpit?

The issue of effective human control in the cockpit burst into the public view following two tragic aviation accidents involving the Boeing Company, the world's most trusted aerospace corporation. Boeing's stock plummeted, prompting the U.S. president to publicly state that the company was too important to go bankrupt. What was the root cause of the tragedies and near bankruptcy? Over-automation and human deskilling of the pilots of the venerable Boeing 737. While details

of the incident are too complex and voluminous to review here, suffice it to say that Boeing pursued software solutions to solve aeronautical engineering issues relating to the positioning of the engines on the wings. The software proved too complex and too automated, and the pilots of two planes were unable to overcome a computer-generated dive, resulting in the loss of all souls onboard both aircraft.

Compounding the problem was Boeing's denial of the root problem after the first accident, trusting in the advanced software, until the second accident made the evidence incontrovertible. Thus, the question confronts us: For civilian airlines, how much automation and how much human skill? Where is the balance? And, with AI advances, there may be more pressure to replace pilots, but when one considers the possibility of hacking and AI hallucinations, a go-slower approach to preserve human-machine resilience in the cockpit seems the right path.

Similar questions confront the military. A debate has raged in the Air Force, Army, Navy and Marine Corps about the balance between manned and unmanned cockpits of the future. While we will have both going forward, the question of cybersecurity should give cause for pause. In the later years of the Afghanistan conflict, an advanced U.S. surveillance drone was downed by a relatively primitive 'spoofing' or hacking of the GPS signal.[xxxvii] No doubt the Air Force has hardened drone defenses against such primitive hacking, but hackers can upgrade their technical tool set, too, ad infinitum. One need only consider: Would a human reconnaissance pilot have allowed his/her aircraft to turn west and head over Iranian airspace? Not under any circumstances. Thus, we should ask: Is it wise to increasingly turn over surveillance to systems that, if we lose cyber superiority, we lose the surveillance fleet?

These are complex problems that will only grow more portentous as DOD works to integrate AI and ever-more robotic platforms into our frontline forces. Given the momentum phenomenon discussed before, it is urgent to slow the deployment of AI and automation and preserve more naturally skilled humans in more cockpits, while the longer-term implications of emerging technologies become clearer.

› Nuclear Forces Risk Curve in Age of AI/ Cyberwar

When the triad of nuclear deterrence (air, sea and land-based nuclear delivery systems) was first constructed, cyber conflict did not exist. Has the emergence of rapidly accelerating AI, combined with cyber conflict, shifted the risk curves, such that trust and resilience, and fundamental human control, may be at risk? As mentioned above, the Russians have already over-automated a nuclear-armed submarine. We need to carefully consider the arguments against ever making the next nuclear bomber unmanned. Again, a similar refrain should come to mind: Why take the bomber pilot out of the cockpit? Is this argument one that again privileges automation and faster decision-making? The decision to start a nuclear war to destroy an opposing country should be conducted at human speed with a premium on human trust (two or three persons in the cockpit with verified orders from the White House—not a hackable computer).

Given the increasing reality of cyberwar and the possibility that our nation could lose cyber superiority, might it be time to consider a radical possibility: that all the nuclear deterrent forces should be human guided, that nuclear missiles will be limited in range such that only a pilot or submarine captain who navigated within the maximum stand-off range could launch an attack of such missiles? What is the risk-trade calculus if an enemy hacked a single nuclear missile and redirected its course away from an enemy state to that of an ally? Again, humans in the cockpit or humans at the helm of a submarine, with limited-range nuclear missiles, reduce the surface area of attack, increase resilience, make silent failures unlikely, and provide a cognitive system that needs fewer patches and expensive cyber software upgrades. Minimally, DOD should resist efforts by OODA-Loop philosophers to integrate nuclear command and

control into larger networks, especially if powered by increasingly capable AI.

Yet, this latter scenario seems to be under consideration in the latest technostrategic 2023 document coming out of the Pentagon, proposing the Joint All Domain Command and Control (JADC2) concept.[xxxviii] While I am not privy to highly classified nuclear deterrent discussions, it seems now is the time to begin IDD control talks, especially regarding nuclear-armed devices, with the Chinese and the Russians, rather than another attempt to make American retaliatory strike capability even faster by integrating these doomsday weapons into JADC2 digital architecture.

› Military SCADA

The threat of cyber insecurity regarding military Supervisory Control and Data Acquisition (SCADA) systems poses an existential risk for the nation's security and the lives of servicemembers on the front line. To achieve resilience, humans must be put back in the control room at the breaker panel with the skill and knowledge needed to effect basic repairs. As the military increases reliance on AI, automation and robotics, what emerges is the proverbial black box of declining human understanding and the ability to explain. Already, explainability is a challenge of some significance for AI scientists, but for such a situation to develop in the military may be tantamount to dereliction of duty. Without questioning the assumptions of speed and efficiency, we risk trust and resilience as the influence of private tech and defense contractors increases, and simultaneously the human skill and understanding of military officers and enlisted personnel declines. Is this fated? Or once again a default choice?

As a former chief nuclear engineer, I knew my technicians, all sworn to defend our Constitution and not conflicted by corporate loyalties, could understand and mitigate failures on most ships' critical systems. Now, on modern ships, the vulnerability of ship SCADA-like systems and those ashore, dependent on millions of IDDs, causes me to pause.

While it is impracticable to reestablish the human skill base and knowledge to provide backup operations for many ship systems, DOD must carefully consider the reinstallation of basic control systems to enable the ship or base to provide basic SCADA-like services, such as keeping the water running, the lights on, minimal propulsion and the ability to return to base—or what sailors on my ship referred to as the "Get-Home Box," a bypass to the advanced electronics allowing sailors to drive the ship with basic electro-mechanical signals to the steering system and auxiliary propulsion. Although anachronistic sounding, in the face of loss of cyber superiority, these human-centric back up measures are becoming more logical.[xxxix]

› One Cloud or Many?

In the 1950s, the U.S. Navy planned to solve the complexity of digital networks with a single ship carrying a large mainframe computer to broadcast to the whole squadron. This centralized concept, analogous to today's Cloud, was considered too vulnerable and replaced with distributed, independent computers on every ship, capable of operating in a completely stand-alone mode. Similarly now, a centralized Cloud computing solution for military operations produces the same obvious risk. The additional risk is as mentioned above: Who is essentially in control? Cloud technicians in Silicon Valley conflicted by corporate loyalties; the programmers of AI, which now controls the Cloud of military data; or at some minimum level, is control preserved under uniformed service members sworn to protect the Constitution?

Many studies indicate that warfare is evolving away from larger military platforms toward smaller and more numerous "swarms" to create more survivability and lethality. If we believe the computer and tactical kinetic battlefields have something in common, then might we need to reconsider by analogy that when seeking resilient, survivable computational storage capability, smaller and more numerous Clouds are better than bigger and singular? At one point, DOD was running almost 5,000 quasi-independent networks,[xl] which might be considered a swarm of networks.

Hopefully, somewhere in a deeply classified computer-storage wargame, senior decision-makers

are considering the worst-case scenario of all-out cyberwar and the benefits or costs of swarm storage and computation strategy.

› AI/Cyber Age Conundrum

In the quest for speed and efficiency, do we risk compromising the trust and resilience of the U.S. military, as powerful AI combines with the tools of cyber conflict? Sometimes the aggressive pursuit of increased decision speed, as in OODA-Loop orthodoxy, is well justified. In other cases, several crucial systems were digitized and automated before the age of AI-powered cyberconflict came into focus, precipitating a tradeoff for speed and efficiency over trust and resilience. Keeping more humans in the loop and retraining them to regain lost skills may be a crucial strategy to improve the security of the nation and individual states in this era of AI cyberpower, characterized by its opaque calculations. The decision to shape, narrow or slow the trajectory of AI and automation and to preserve a modicum of human knowledge, skill and abilities will be unpopular in the defense industry. But asking hard questions of technical elites and reconsidering tradeoffs has a long history in our nation, and we are well justified in invoking this prudent tradition at the dawn of the AI/Cyber Age.

And we need to do so urgently. In September, according to The Hill, "Deputy Secretary of Defense Kathleen Hicks … touted a new initiative designed to create thousands of [autonomous] weapons systems powered by artificial intelligence, saying it will mark a 'game-changing shift' in defense and security as Washington looks to curtail China's growing influence across the world. … [T]he new initiative, called Replicator, is part of a concentrated push at the Pentagon to accelerate cultural and technological change and gain a 'military advantage faster' over competitors. … [T]he Pentagon would work closely with the defense industry to field thousands of autonomous weapons and security systems across all domains in 18-24 months."[xli]

Moreover, given the high-tech, digital-savvy Israeli Defense Force and our own CIA/NSA were completely surprised by a low-tech, human-centric attack of strategic proportions on October 7 (the 50th anniversary of the 1973 Yom Kippur war), is more digital tech the answer? The rush to digitize and replace humans with intelligent machines may be ill-advised at this time of strategic uncertainty, and a thoughtful, slower approach seems in order. Political leaders need to start asking the hard questions, now.

---

[i] Russell, Stuart, *Human Compatible: Artificial Intelligence and the Problem of Control*, New York: Viking-Random House, 2019, pp. 255-56, extended quote here: "The solution to this problem [preserving human autonomy] seems to be cultural, not technical. We will need a cultural movement to reshape our ideals and preferences towards autonomy, agency, and ability and away from self-indulgence and dependency—if you like, a modern, cultural version of ancient Sparta's military ethos."

[ii] Kissinger, Henry, Eric Schmidt and Daniel Huttenlocher, "ChatGPT Heralds an Intellectual Revolution: Generative artificial intelligence presents a philosophical and practical challenge on a scale not experienced since the start of the Enlightenment," *Wall Street Journal*, 24 February 2023, accessed here: https://www.wsj.com/articles/chatgpt-heralds-an-intellectual-revolution-enlightenment-artificial-intelligence-homo-technicus-technology-cognition-morality-philosophy-774331c6

[iii] The series of war games was funded by the Office of Secretary of Defense, Office of Rapid Fielding. Project leadership was shared between Dr. Peter Singer of Brookings and the NOETIC Corporation.

[iv] Hagerott, Mark, "Lethal Autonomous Weapons Systems: Offering a Framework and Some Suggestions," presented at the Geneva convening of the 2014 Convention on Certain Conventional Weapons (CCW), link to brief here: https://docs-library.unoda.org/Convention_on_Certain_Conventional_Weapons_-_Informal_Meeting_of_Experts_(2014)/Hagerott_LAWS_military_2014.pdf

[v] Technology has been shown to be susceptible to what is called "momentum" or "technological lock-in," wherein early decisions may gain a kind of inertia, and later efforts to redirect technology's trajectory become all but impossible often with negative consequences for later generations. See Thomas Hughes' study of electrical power networks and Paul David's study of the QWERTY keyboard.

[vi] The expanding realm of cyber insecurity is penetrating an increasing number of activities, from email servers, government data bases, banks and critical infrastructure to now frontline weapons. See Heckmann, Laura, "Trustworthy Tech: Air Force Research Lab Looking at Uncertainties with Electronics," National Defense, August 2023, pp. 28-30.

[vii] The advocates for accelerating the acquisition of unmanned systems are many, but their acknowledgement of the potentially high costs of ensuring cyber security could be more candid. If the lifetime costs of these systems included never ending cybersecurity contracts, the argument to automate might be less compelling. In the first war game of the OSD series, for example, one senior level robot company executive exemplified this problem. When pressed about the cybersecurity of his company's unmanned systems during an off-the-record meeting, the executive deferred cyber insecurity to software companies. He appeared to take little ownership of the potentially massive problem and offered that the "banks would be the first to solve the problem."

[viii] The definition used in this essay is a simplification of many longer attempts at explanation. A particularly thoughtful essay on the subject is by Joe Nye, Jr., See "Cyber Power" at: https://projects.csail.mit.edu/ecir/wiki/images/d/da/Nye_Cyber_Powe1.pdf

[ix] "Silent failures" are considered by some experts in the field to be the worst kind, potentially the most damaging, since you don't know they occur. With a human in the loop, especially on physical-kinetic type platforms, a human is on scene and can more quickly identify if the platform or system is failing to follow the assigned tasks. For more on silent failures, see Dan Geer, 26 May 2013, interview, "The most serious attackers will probably get in no matter what you do. At this point, the design principal, if you're a security person working inside a firm, is not failures, but no silent failures." Accessed on 31 July 2013: http://newsle.com/article/0/77585703/

[x] Martin Libicki of the RAND Corporation has written extensively on the problem of cyber attribution and the differences with traditional theories of deterrence, which relied on more certain knowledge of a potential adversary's military capabilities than might be possible in the case of cyber.

[xi] Geer, Dan, 26 May 2013, interview with Newsle.com. Geer, a leading CIA executive, has noted that the complexity of our electronic netted systems may be the biggest challenge going forward, even before accounting for the determined attacks of a cyber rival.

[xii] Luckenbaugh, Josh, "Algorithmic Warfare: DARPA Host Workshops to Develop 'Trustworthy AI,'" National Defense, June 2023, pg. 7

[xiii] Heckmann, Laura, "Trustworthy Tech: Air Force Research Lab Looking at Uncertainties with Electronics," National Defense, August 2023, pp. 28-29.

[xiv] CJCS U.S. Army Gen. Martin E. Dempsey's speech at Brookings, 27 June 2013, warned of the growing scope and dynamism of cyber warfare, which will only grow in significance.

[xv] John Boyd exerted his greatest influence in the Air Force. In the Navy, RADM Wayne Meyer and VADM Art Cebrowski were the strongest advocates for automation and speed. Meyer was the lead architect of the AEGIS combat system, which was capable of autonomous/automatic weapon assignments and engagement (though the computer system was carried aboard a manned platform). Cebrowski is credited with developing the concept of Network-Centric Warfare, the idea that war would be fought between networks of high-speed computers and communications links. Interestingly, a subculture of the Navy, Rickover's nuclear engineers, have been more reluctant to embrace ubiquitous networks, OODA-Loop speed and Network-Centric Warfare as have other communities. Submariners generally tend to value the independence of command and distributed, non-netted warfare. See article by Michael Melia, "Michael Connor, Navy Vice Admiral, Calls For Submarine Commanders' Autonomy," accessed 17 July 2013: http://www.huffingtonpost.com/2013/07/17/michael-connor-navy_n_3612217.html

[xvi] We have seen this momentum effect in DOD's R&D/acquisition system. See Donald McKenzie, *Inventing Accuracy: An Historical Sociology of Nuclear Missile Guidance*, Cambridge: MIT Press, 1990.

[xvii] The literature regarding Col. John Boyd is extensive. For a lengthy treatment, see Robert Coram, *Boyd: The Fighter Pilot Who Changed the Art of War,* New York: Hachette Book Group, 2010.

[xviii] Cranberry, Sean, Special Report, Part 1 of 7: "Joint All-Domain Command, Control A Journey, Not a Destination," National Defense, July 2023.

[xix] Cyber resilience is becoming something of a watchword. Perhaps the most compelling definition is: "An operationally resilient service is a service that can meet its mission under times of disruption or stress and can return to normalcy when the disruption or stress is eliminated," on pg.1 in "Measuring Operational Resilience," Software Engineering Institute, Carnegie Mellon University. But we should ponder: If a system was even temporarily hacked, or trust broken, the issue of resilience may become moot, because if the ship were to run aground or the aircraft crash, even should the computer programmers regain cyber control, the system is irreparably damaged even after a brief loss of cybersecurity. In these cases, we need to maintain and reinsert human operators and human skill … to stay off the rocks, so to speak.

[xx] Milley, Gen. Mark, Chairman of the Joint Chiefs. "Strategic Inflection Point," Joint Forces Quarterly, 110, 3rd Quarter, 2023, pg. 6.

[xxi] National Science Foundation, "Artificial Intelligence (AI) Research Institutes, Program Solicitation, NSF 23-610, Theme 3," accessed 17 August 2023: https://www.nsf.gov/pubs/2023/nsf23610/nsf23610.htm?WT.mc_ev=click&WT.mc_id=&utm_medium=email&utm_source=govdelivery

[xxii] In the field of technology studies, a key concept is technological momentum, wherein technological programs build up financial and social momentum, and sometimes propel themselves well beyond the size and scope that could be justified by dispassionate analysis. For discussion of this concept, see Thomas Hughes, *Networks of Power*, and Donald MacKenzie, *Inventing Accuracy*, the latter providing evidence of nuclear missile improvements, which in the later Cold War gained a momentum provided by R&D/acquisition programs that exceeded most reasonable requirements when viewed in hindsight.

[xxiii] Peterson, Dale. "People are Not THE Answer," Digital Bond, 1 May 2013. See: http://www.digitalbond.com/blog/2013/05/01/people-are-not-the-answer/

[xxiv] DOD has announced that CYBERCOM will expand significantly, both with uniform personnel and contractors in an effort to provide better and more numerous experts to run cyber defense.

[xxv] The failure of human technicians to audibly monitor the centrifuges at the Natanz nuclear site was a key element in the STUXNET attack. Alert human operators would have heard the effects of the attack and could have stopped it, but they remained fixated on their digital and hacked instruments and were oblivious of the attack underway. See Mark Hagerott, "Stuxnet and the Vital Role of Critical Infrastructure Engineers and Operators," International Journal of Critical Infrastructure Protection, (2014), v7, pp 244-246.

[xxvi] While the veracity of such articles may be in doubt, it is interesting to note that reports out of Russia indicate that in the wake of the Snowden disclosures, the Russians are reverting some critical communications to manual, human-centric typewriters and human-readable paper, unmediated by any electronic devices. This was a similar measure adopted by Osam Bin Laden and Al Qaida, when they reverted to human messengers for their communications. The American reader's reaction is most likely that we are better than the Russians and Al Qaida. But in the face of a determined cyberpower or following treacherous actions by an insider, such measures should at least be trained and planned as a contribution to overall military operational resilience.

[xxvii] See discussions of "resilience," especially recent studies conducted by Carnegie Mellon University, Software Engineering Institute.

[xxviii] The Resilient Navigation and Timing Foundation has worked assiduously to call attention to the cyber vulnerabilities and existential threat to navigation in our current GPS dependent condition. See: https://rntfnd.org/

[xxix] The tendency of billions of humans to prefer faster, clipped, easier communication via Twitter, texting and hyperlinks is now showing potential signs of deskilling cognitive functioning. See Nicholas Carr, *The Shallows: What the Internet is Doing to Our Brains*, New York: Norton, 2010. There is growing evidence that naval officers are falling into the same deskilling pattern to destructive effect. See USS Guardian grounding report: http://www.cpf.navy.mil/foia/reading-room

[xxx] Haney, ADM Cecil D. Report by Commander, Pacific Fleet, 22 May 2013, http://www.cpf.navy.mil/foia/reading-room/

[xxxi] The ability to hack any number of medical devices has now been established, and the FDA is moving quickly to control their proliferation. See Sun, Lena H. and Dennis, Brady, "FDA, facing cybersecurity threats, tightens medical-device standards," Washington Post, 13 June 2013 accessed on 31 July 2013 here: http://articles.washingtonpost.com/2013-06-13/national/39937799_1_passwords-medical-devices-cybersecurity

[xxxii] Study is being led by the Office of the Undersecretary of Defense for Acquisition and Sustainment, Dr. David Restione, Director of DOD Wearable Pilot Program. See Indo-Pac Conference, 2022, pg 11, accessed https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2022/post/agenda.pdf

[xxxiii] North Dakota was the first state in the Union to require cyber PK-20 education standards and also to propose a new Digital Cyber Land Grant Act. See Anderson, Tim, CSG Midwest Newsletter, 28 June 2023, link here: https://csgmidwest.org/2023/06/28/north-dakota-is-first-u-s-state-to-require-cybersecurity-instruction-in-k-12-schools/. The North Dakota University System the first to propose a new Digital Cyber Land Grand Act. See Hagerott, Mark, "Time for a Digital Cyber Land Grant University System," Issues in Science and Technology, Winter 2020, xxxiii, accessed here: https://issues.org/time-for-a-digital-cyber-land-grant-system/

[xxxiv] Morison, Elting, *Men, Machines and Modern Times*, Cambridge: MIT Press, 1966.

[xxxv] "Pause Giant AI Experiments: An Open Letter," 22 March 2023, See: https://futureoflife.org/open-letter/pause-giant-ai-experiments/

[xxxvi] Huet, Natalie, "What is Russia's Poseidon nuclear drone and could it wipe out the UK in a radioactive tsunami?" Euro News, 5 May 2022, accessed here: https://www.euronews.com/next/2022/05/04/what-is-russia-s-poseidon-nuclear-drone-and-could-it-wipe-out-the-uk-in-a-radioactive-tsun. See also, Kaur, Silky, "One nuclear-armed Poseidon torpedo could decimate a coastal city. Russia wants 30 of them," Bulletin of the Atomic Scientists, June 14, 2023, accessed here: https://thebulletin.org/2023/06/one-nuclear-armed-poseidon-torpedo-could-decimate-a-coastal-city-russia-wants-30-of-them/

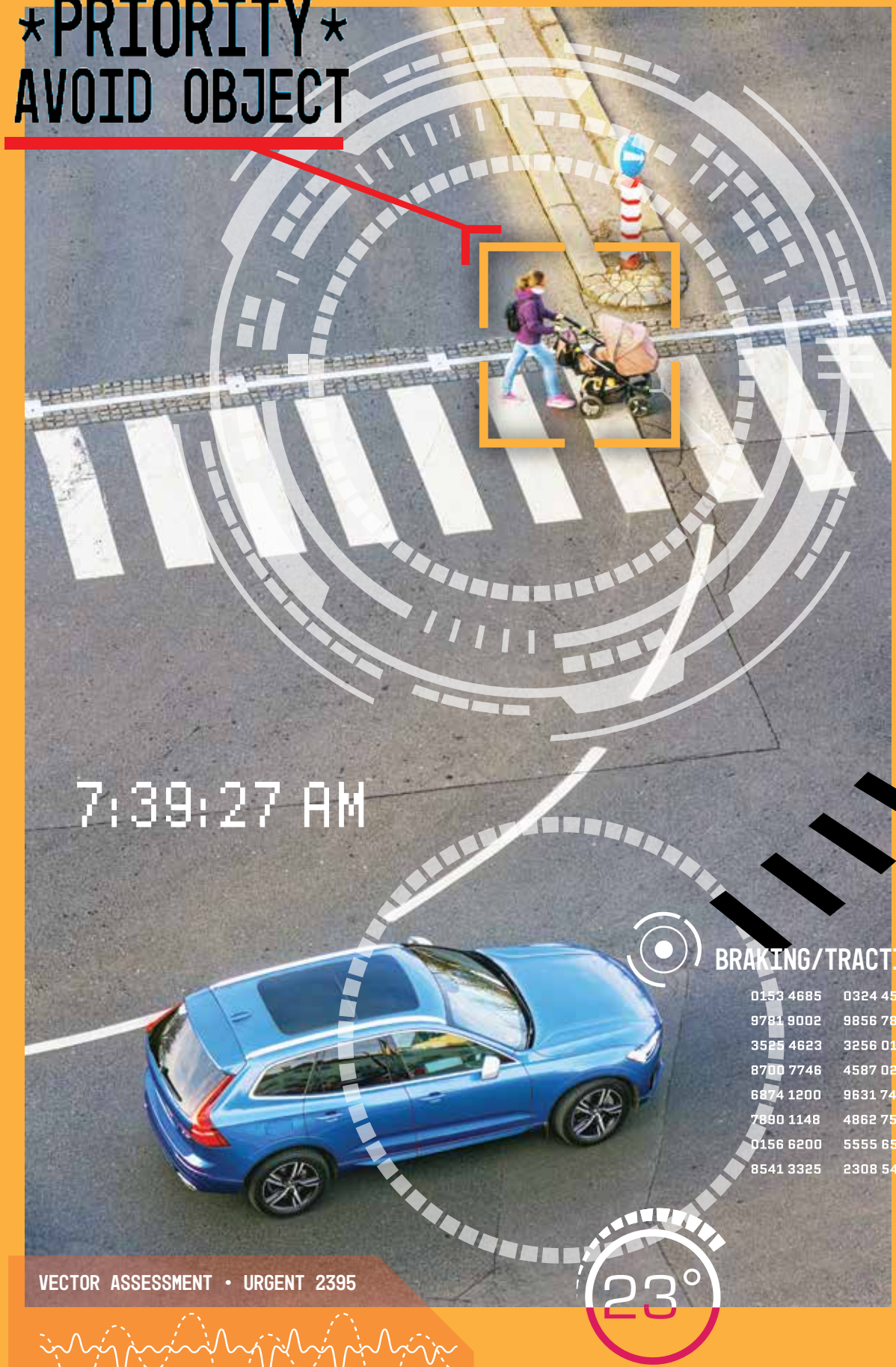
[xxxvii] Shane, Scott and David Sanger. "Drone Crash in Iran Reveals Secret U.S. Surveillance Effort," New York Times. 7 December 2011.

[xxxviii] Cranberry, "Joint All-Domain Command, Control A Journey, Not a Destination."

[xxxix] Fahmida Y. Rashid, "Internet Security: Putting the Human Back in the Loop," 19 April 2012, See: http://securitywatch.pcmag.com/security/296834-internet-security-put-humans-back-in-the-loop

[xl] Dempsey, General Martin. Brookings Institution, 27 June 2013.

[xli] Dress, Brad, "Pentagon's New AI Drone Initiative Seeks Game-Changing Shift to Global Defense," 9 September 2023, The Hill.


PRIORITY
AVOID OBJECT
7:39:27 AM
BRAKING/TRACTION
0153 4685 0324 4578
9781 9002 9856 7854
3525 4623 3256 0125
8700 7746 4587 0258
6874 1200 9631 7459
7890 1148 4862 7592
0156 6200 5555 6587
8541 3325 2308 5471
23°
VECTOR ASSESSMENT • URGENT 2395

A Duty to Use Autonomous Vehicles?

When Safer Than Average Human Drivers

Dennis Cooley, PhD
Professor of Philosophy and Ethics, NDSU

Public opinions on technology run the gamut from the best, most exciting, progressive thing humans create to improve their lives and the world around them, to posing a threat to our species' existence.

There is technology deserving that approbation, such as more effective cancer treatments, DNA testing kits and GPS guided tractors. We would not be too far out of line to say it is almost miraculous. However, other innovations fall on the opposite end of public opinion: Any device or application relying on artificial intelligence (AI) beckons demise via Arnold Schwarzenegger's Terminator cyborg assassin.

Certainly, there are dangers. "We are entering a time when technological evolution creates two new realms of socio-economic-technical activity: the Integrated Realm and the Machine-Robotic Realm," Mark Hagerott, PhD, the Chancellor of the North Dakota University System and a technology expert, told my NDSU students in a talk this spring. (Please refer to the illustration on page 42.) "The capacity of AI/autonomous machines (robots) and cyberspace (the metaverse) to create wealth and power is enormous. But this technology has near unlimited power to impinge on the sanctity of human space, a possibility once protected against by the physical frontier of technology."

Yet in the real world, digital technologies are proving quite valuable tools for human use. For example. vehicles that reduce accidents by helping drivers stay in their lanes and otherwise make driving safer when drivers aren't paying sufficient attention. People are glued to their phones and other devices, and GPS has become essential in some lives. As a result, attempts to reduce AI's encroachment into research and the marketplace are likely to fail. AI is too convenient, helpful and profitable.

On major question is whether the most beneficial AI technologies, such as using self-driving vehicles (which are not necessarily electrical), should become a duty rather than merely an option. If there is a duty to use them, then is there a corresponding obligation to create them?

Regarding self-driving vehicles, a 2022 study predicted that 94 percent of long-haul commercial trucking can be done by AI provided that the technology improves to handle all weather conditions.[i] If the study's authors are correct, high turnover rates and the current shortage of 80,000 drivers can be mitigated if not eliminated. (Of course, the potential for significant job losses with AI adoption is an important and controversial topic, but beyond the focus of this article.) And let us not forget the potential safety improvements through devices that never tire or become distracted. If self-driving

Illustration / Jerry Anderson

An example of human lives with too little "Human-Natural" and too much "Machine-Robotic" and "Cyber Space," as illustrated in the 2008 film "Wall·E."

technology can work for truck driving, then there is no reason to believe it can't work the same for long-haul car trips, or as a designated operator for the elderly, or for those with substance abuse or medical problems, or others at risk or risky drivers.

Granted that adopting technology is enticing, especially when it eliminates significant problems, ethical worries arise. A more significant one is technology's impact on human beings in their lived environments. In particular, through AI, eroding human freedom in making decisions, which has negative impacts on autonomy, moral agency, and people's abilities to be and do what they should. More simply, by replacing human decision-making with that of a program or machine, we make people less able to make autonomous choices for themselves and therefore infantilize them rather than empower them. If this is part of a slippery slope argument, then we could very well end up producing the rotund, apathetic human survivors encountered in the film "Wall·E."

Argument for a Mandate

Freedom and free will are two of the most valued human powers because they are essential to us being moral agents in the first place. To limit either, therefore, has to be justified with far stronger evidence than merely defending why people are entitled to exercise them, and then letting the social marketplace sort it out. To propose limiting freedom and free will, such as having a duty to buy and use a self-driving car, demands even more justification. This is in fact an extraordinary claim requiring extraordinary evidence.

Deciding if there is a duty to do something, furthermore, requires a higher standard than merely proving that an action is morally permissible or right. Duties entail that failure to perform them is automatically forbidden and wrong, unlike an act being morally right, which might mean it is one of many morally right actions. The difference here is between it being permissible to buy a self-driving car versus a standard vehicle, and the moral (not legal) mandate that only a self-driving car will fulfill one's duty.

Many moral factors are at stake in deciding if technology is permissible, much less obligatory. In engineering, there are five ethical factors that help decide when a risk is morally acceptable, which also can address when technology is permissible or required:

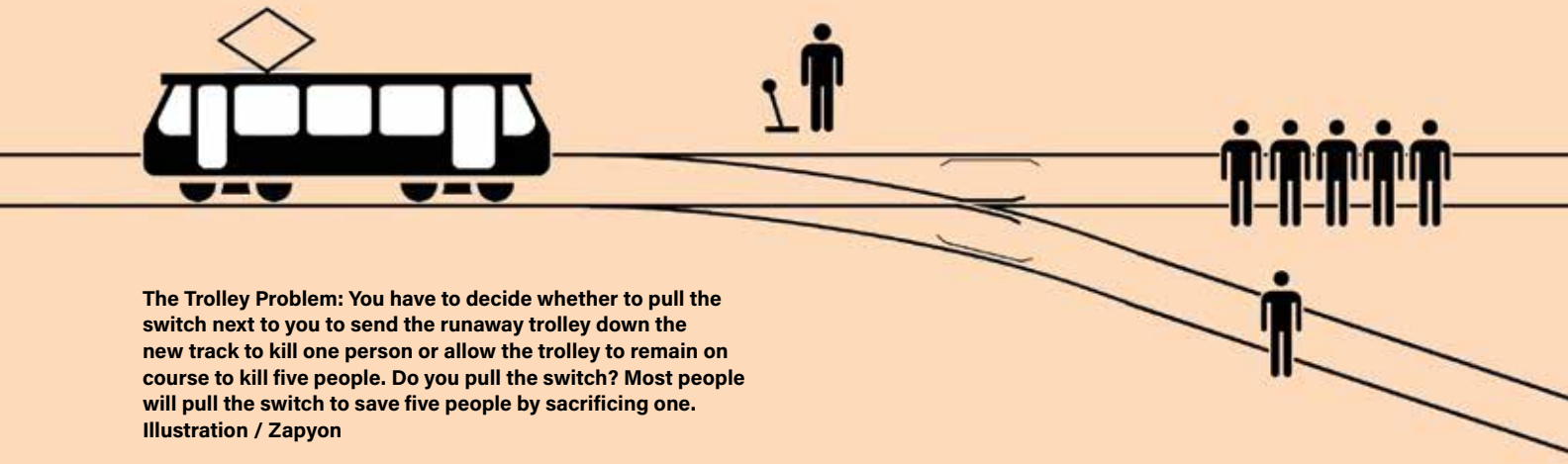
A The degree of informed consent with the risk,

B The degree to which the risk is voluntarily accepted,

C The degree to which the benefits of a risky activity weigh up against the disadvantages and risks,

D The availability of alternatives with a lower risk, and

E The degree to which risks and advantages are justly distributed.[ii]



The Trolley Problem: You have to decide whether to pull the switch next to you to send the runaway trolley down the new track to kill one person or allow the trolley to remain on course to kill five people. Do you pull the switch? Most people will pull the switch to save five people by sacrificing one. Illustration / Zapyon

For the first two, above, freedom and free will entail that if people understand the risk to themselves and still decides to engage, then their decision should be deferred to. They have the right to make that decision and also the responsibility for the consequences, good or bad. On the other hand, imposing hazard on others without their knowledge or consent is generally impermissible because this does not respect their free-will agency. The third criterion is merely a cost-benefit analysis that the technology has to be worth the cost, whereas the fourth factor states the common-sense view that any option, which gets us where we want to go without as much risk to self or others as the other alternatives open to us at that time, is the only rational option to pursue. The final factor concerns justice: We should not impose greater risk on more vulnerable members and groups of our population, especially if the rewards are not sufficiently shared with them. The unfairness becomes greater if the vulnerable are the only ones to bear the costs, while the privileged receive all the benefits.

Given these moral factors, when can a moral agent be obligated to use technology? When the risk of not doing so is so great that it must be mitigated or eliminated. More precisely, the person must use the technology if all of the following five conditions are met. Firstly, if adopting the technology significantly lowers the risks involuntarily imposed on others and in which the risks for severe harm to either the agent doing the action or those affected by it are high. The third through fifth requirements are that there isn't a considerably better alternative that achieves the desirable outcomes more efficiently and through which the risks are significantly more equitably distributed, *while* at the same time the technology does not burden the agent to an excessive degree

The Technology

Most AI implementation begins with the basics of rational decision-making for actions: Values and other relevant factors are the construction material, and principles are the tools to put the material together in various, approved ways. Once identified through theoretical problems, such as Trolley Problems, some decision procedures, with carefully delineated steps to build a solution, are programmed. The goal here is for AI to identify the rules and relevant information, and then manipulate them successfully to show at least a nodding acquaintance with how humans make decisions and work.

Of course, at the moment, the reliability of driverless cars is far below what is required to make them a less risky alternative to normal driving on average. That means that before autonomous driving systems could replace all human drivers, they must be able to reduce travel risks to a level significantly below that of the average driver for that population demographic. Moreover, this technology would have to be acceptable to the general community, based on the criteria above.

One general standard is that the technology has to be safer than the average human driver, according to Michael A. Nees.[iii] It is hard to understand what this requirement actually means in practice because most people believe themselves to a better driver than the average. So, is the standard really about being better than the average human driver or something much higher?

That seems to be the case in a survey on driverless cars, conducted by the City, University of London (CITU).[iv] Sixty-one percent of respondents said the technology and cars would have to be much safer than the safest human driver or the standard increased to never causing a serious collision. People don't trust self-driving technology: Only 18 percent said they were comfortable with autonomous cars on the road as long as they are as safe as the average human driver. It, therefore, might be best to concentrate on drivers who are more at risk than the average driver and merely use the average driver as our measure for self-driving car technology's permissibility.

AI in motor vehicles makes sense when it reduces accidents by augmenting human driving, such as the stay-in-one's-lane technology. Depending on the situation, it may or may not be essential to adopt these safety features. The case for a duty grows stronger when the technology becomes the only reasonable travel option for a driver with medical or other issues that greatly increase risk on the road. Drunk drivers, many elderly and people with some medical conditions move the risk from that faced in normal driving conditions with normal drivers to a far higher qualitative and quantitative degree.

Furthermore, there is an opportunity benefit to those who cannot drive because of some medical conditions or other pressing problem: They have the freedom to have a car dedicated to their needs and decisions rather than having to rely upon public transport or other people's schedules. This ability to do what one needs or wants to do is liberating by giving people control over some of the basic activities taken for granted in a car-centric society, especially in rural states where there is no public transportation, or it can be an imposition to ask neighbors and family.

AI's Threat

"Man is born free and everywhere he is in chains," wrote Jean-Jacques Rousseau in *The Social Contract,*[v] aptly capturing humanity's current condition, especially in the industrialized world with its dependence on its technology. Most of us feel incomplete, for instance, without constantly checking our smartphones to see if there is an email, text or some oddly interesting new posting on social media. We are lost without access to the digital world because it is essential to being able to function and thrive in our technology dependent society.

Technology is supposed to liberate us from repetitive drudgery to do more interesting things, but it can end up doing the opposite. Technological determinism, a sociological term for the fatalistic surrender to technology, binds us to a world that is determined by technology rather than our lives being under our meaningful control. Something like this was predicted by Martin Heidegger in "The Question Concerning Technology." There he writes that while technology has no inherent moral value, the way humans approach crafting the world in which they live as a response to technology makes it conditionally good or bad.[vi] The problem, he says is the way humans have myopically adopted technological thinking—calculative reasoning—as the only form of thought. This in turn has caused us to begin seeing people and all things in the universe as mechanical objects rather than for what they truly are.

This mechanistic, mental framework leads us into inauthentic instead of the authentic existence, which everyone naturally seeks. Instead of having an existence filled with wonder and amazement, the result is the "I" of the individual grasping his true Being is sacrificed to the "they" mentality in which the focus is on objects outside of who we really are. In other words, we fatalistically allow technology to rob us of our freedom to make our lives meaningful, rather than using it for the tool it was intended to be to improve lives.

Humans, moreover, are social animals, who learn what it is to be human through interactions with others and making their own decisions about those interactions. We could say that people are the result of evolutionary

adaptations in which those better able to compete and collaborate in an environment tended to survive and reproduce, thereby passing their genes on to future generations. Part of what made our ancestors better contenders was the ability to make the right decision in the situations they encountered, especially risky ones. Better choosers of the right thing survived, while slower ones became lunch for predators or otherwise had truncated lives. So, it could be reasonably claimed that to be human requires that we make choices for ourselves as individuals, without forgetting that those selections are being performed in collaboration with the other people in our society. Accordingly, although there might be explicit interactions and planning, we can manage driving a vehicle in crowded traffic, walk in a crowd or socially interact with strangers without mishap. To be human, therefore, is to be constantly engaged in decision-making with or without interacting with others.

Self-driving cars and other technology can diminish the quality and quantity of this constant choice-making and eventually enslave people depending upon what the technology actually does. The Society of Automotive Engineers (SAE) makes a distinction between levels of automation for self-driving vehicles, which usefully shows when the technology is operating as the tool to assist and when it can lead to infantilizing humans:

1 Cars with some driver's assistance, such as cruise control and lane change monitoring and warnings.

2 Car with advanced cruise control or an autopilot system that can take safety actions, such as braking.

3 Cars requiring a human driver but able to perform some safety critical functions, such as steering and braking at the same time, in certain conditions.

4 Cars capable of self-driving most of the time without input from the human driver, but which might be programmed not to drive in unmapped areas in severe weather.

5 Cars with full automation in all conditions.

Tesla vehicles come equipped with the most advanced autonomous hardware and software, including enormous processing power, precise GPS, multiple cameras providing a 360-degree view, ultrasonic sensors and now radar to help navigate, especially in bad weather. According to CEO Elon Musk, Teslas configured with Hardware 3, which was first released in 2019, will have full self-driving capability surpassing human safety levels. Hardware 4, which upgrades the onboard computer and sensors, was released for new vehicles in early 2023. However, Tesla's Autopilot (not yet fully autonomous) has been involved in 736 crashes since 2019, including 17 deaths. Many improvements need to be made before fully autonomous driving vehicles can be deployed safely. Photograph / Wikimedia

All technology performing at Levels 4 and 5 can over time—with enough reinforcement rewarding AI takeover of decision-making that achieves whatever is desired—render human beings no longer able to make their own decisions or make far less autonomous ones when the situation is complicated and the relevant information hard to discern. Instead of being free with their free-will faculty working appropriately, people become mere automatons run by the technology that was designed to enhance their existence. They surrender to technological determinism or worse, the Machine-Robotic Realm in which robots and AI make decisions for themselves and humanity. Elon Musk said that AI is an existential threat to human civilization,[vii] and at Levels 4 and 5, that might actually be the case. The general issue is that AI machines could destroy humanity merely because AI's goal is given priority over everything else, including humanity's existence as organisms and moral agents living in their natural and social environments.

Integrated Solution

In response, Hagerott contends that we as a species using technology in our natural and artificial environments have an obligation to reject technological determinism and the Machine-Robotic Realm.[viii] He argues that we should adopt the Integrated Realm framework, which establishes ethics, policies and law that preserve human command of machines. If we take respecting persons with freedom, free will and moral agency seriously, the Integrated Realm is morally required as the only realm that recognizes what human beings are and how they operate in the real world.

In San Francisco, a self-driving car operated by Cruise, owned by General Motors, ran over a woman after she was knocked in front of it by a hit-and-run driver. The Cruise AV severely injured the pedestrian, and firefighters arrived to find her pinned underneath the vehicle. Firefighters contacted the Cruise control center to make sure the vehicle was securely stopped and then used heavy rescue tools to lift it and pull the woman out, fire department officials said in a press release. In August, California authorities expanded driverless taxi services in San Francisco, giving the go-ahead for Waymo and Cruise operators to compete with ride-share services and cabs. The California Public Utilities Commission voted to let Waymo, a unit of Google-parent Alphabet, and Cruise essentially run 24-hour robotaxi services in the city. Photograph / San Francisco Fire Department.

Besides being free and possessing free will, our consciousness is non-computational as Roger Penrose argues in *Shadows of the Mind.* That means it's not the computational brain model's orderly systems running orderly programs that can be duplicated in computer language code: "There must be more to human thinking than can ever be achieved by a computer, in the sense that we understand the term 'computer' today."[ix] Consciousness and understanding, according to Penrose, can only be explained by figuring out the connection between the quantum and classical physics of how our brains and their components function. Computers today cannot do this since: "Intelligence cannot be present without understanding. No computer has any awareness of what it does."[x] Making computers more and more powerful, with the ability to evolve their own code or perform innumerable Trolley Problem experiments to determine how human beings react in stressful choice situations, will not lead to AI that is conscious or able to make decisions as humans do.

Ethical codes are unique to human beings and essential in every area of social life. They reflect the values and processes that society uses to govern the existence of and interactions among individual citizens, groups and institutions, partly to keep the society functioning acceptably. Moral codes are merely specialized social ethical codes aimed at making community members act ethically. These codes tend to develop over time, not systematically, but rather as the need alteration is perceived. As novel, unforeseen situations arise, there is a tendency to add process rules on what professionals should be or do for future, morally similar occurrences. If something goes wrong, then great pains are taken to revise the code so that the misstep won't recur.

Human morality/ethics doesn't work the way logic does in computer programs, mostly because human beings are not designed, mechanical systems. Ethics and much human activity require people to understand and engage in human activity critically, creatively and emotionally. We *qua* reasonable, social animals are the products of evolution, socialization and self-directed development; hence, we are more like patchwork creatures in our thought processing than we are finely tuned machines. Reason has a necessary role, we all agree, but emotion/feeling is its essential partner. Moreover, morality inherently incorporates imagination and creativity. When we think about what we should do or be, then we are thinking about worlds that may or may not exist. If they exist, then we ask ourselves if they ought to continue doing so. It took creativity to imagine a world without slavery or one in which women are equal to men, and then to dream how to achieve such result. While using identified moral rules as tools is essential to learning about ethics, something more is necessary to be a moral agent, which AI cannot duplicate.[xi]

Consider the following study on self-driving vehicles that shows the inherent need for humanity and morality in driving: The CITU study, cited above, showed that 91 percent of respondents said that being considerate to other road users (including drivers) is as important as following the formal rules of the road, and 77 percent agreed that drivers sometimes have to use common sense instead of just following the highway code to be able to drive appropriately. What these responses show is that driving and all other human endeavors require imagination. Perhaps less emotionally compelling is that it takes imagination to see when the rules do not apply and come up with an acceptable alternative. Being a moral agent and driver requires us to be good critical and creative thinkers. It requires emotive connection, including empathy and compassion. In conjunction with ethical theory, principles and values, which are a rational part of ethics, we as human, moral agents have a fuller Penrosian understanding of what ethics are and how they work in our decision-making than does any current AI technology. And possibly we will have more than any future AI technology can duplicate.

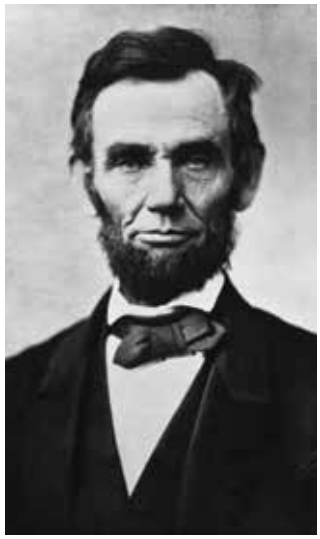
The Integrated Realm framework uses the strengths of people and AI while trying to minimize their weaknesses. Caitlin Deloherty[xii] writes that AI lacks cognition—the human ability to use common sense, intuition, previous experience and learning to make split-second decisions—which makes human beings superior drivers. At the same time, autonomous vehicles have cameras, radar, light detection and ranging sensors that exceed humans' ability to perceive, which help the vehicles navigate better in

foggy or darker driving conditions. So, to get the best of both people and self-driving transport, a balance needs to be struck between when humans have control and in what way, and when to cede control to AI and technology.

SAE's Levels 1 to 3 technology, explained above, is permissible to use and could be morally obligatory in some cases. These mechanical or digital devices augment what people are doing when driving and free some of their attention for more meaningful tasks, such as paying greater attention to the road, driving conditions and other relevant factors not being addressed by the technology. They do not replace human decision-making or the ability of humans

"Freedom is not the right to do what we want, but what we ought."

Abraham Lincoln



to choose wisely. Psychologically as moral agents, we need to develop the brain's executive function through experience, according to "The Adolescent Brain."[xiii] Executive function exerts inhibitory control and includes working memory, which is the ability to keep information and rules in mind while performing mental tasks. Inhibitory control is the ability to halt automatic impulses and focus on the problem at hand. For example, running a meeting in a different way or taking a new, rather than habitual, route to work involves both inhibitory control and working memory. Doing things in new ways requires that people are in charge of identifying what matters, make decisions and plans to carry out their decisions, and then implement them. All of that is driven by executive function.

SAE's levels 1 to 3 technology beneficially alleviate ego depletion, which happens when one's willpower and self-control reduce or exhaust a person's limited pool of mental resources. Of course, that depletion adversely affects executive function. According to one study,[xiv] our executive function starts to perform suboptimally as more and more decisions must be made. As we exhaust our mind with many trivial choices, we lack the mental energy to choose wisely for more important challenges, such as how to handle an unexpected item blowing directly toward one's speeding vehicle. So, not all decision-making is good if it makes us too tired to perform well when that matters more than merely deciding. By allowing humans to focus on the more important choices, AI and levels 1-3 technology are, therefore, morally permissible.

A Moral Mandate?

Self-driving cars and other related technologies become mandatory for different reasons at the five levels. For SAE's automation Levels 1 to 5, it is merely permissible for average drivers, or within the standard deviation of being so, to use the various technologies or not. For the first three levels, it would be a good idea to drive with these technologies in order to keep the drivers' executive function higher and ego depletion lower for more important matters that may arise during a trip. Levels 4 and 5, on the other hand, pose more safety risks than the others if they begin infantilize drivers by making them less able to think quickly, creatively, critically or pragmatically while in the vehicle or in the driver's other life experiences. But as long as these normal drivers retain the skills that make them human and thrive, the technology here is permitted but not required.

Levels 1 to 5 become more likely to be mandatory for those who have impairments that make their driving riskier than for the average driver. How to determine whether there is a duty to use the technology for these groups of drivers depends on five criteria (A through E) identified on page 52, above.

The first three SAE levels of automation could be required for new drivers who need experience to eventually become average enough to no longer need the technological training wheels. Consider that younger or other new drivers are at far greater risk of a motor vehicle crash than others. Males 16

to 19 years of age, for example, are three times more likely to have an accident than female peers. Also, when teenagers drive with other teens or young adult passengers, the chances of an accident increase greatly. Most of the heightened risks stem from inexperience and distractions while driving, according to the 2020 National Automobile Safety Administration report. These drivers need to learn how to drive while using the technology that makes them far safer.

On the other hand, for these drivers, the last two levels would be destructive to their learning how to drive, since they would always basically be passengers, or even worse, in the case of Level 4 made to drive in the worst possible conditions when their skill sets are not up to par. Since we want these drivers to learn to drive and improve their decision-making skills, Levels 4 and 5 cannot be required, unless in very unusual circumstances, such as those below.

SAE's Levels 4 and 5 would be mandatory for only a small slice of the population, including drivers with physical or mental health or other risk-increasing conditions, according to their levels of inability. Someone who is blind would need a Level-5 vehicle. Drivers who have shown repeated refusal not to drink and drive might require a Level-5 vehicle to transport them, whereas those less incorrigible could make do at Level 4 technology. The desirable outcomes achieved by levels 4 and 5 technology enable people with these challenges to have meaningful functionality in a car-dominated society and the ability to make their lives authentic through decision-making, while reducing potential harm to others. Finally, mandating Level 4 and 5 technology for these groups of citizens places the benefits and burdens where they justly should go, at the same time making the world a better place. In other words, it creates a moral, meaningful Integrated Realm.

Future Moral Dilemma?

At this time, there is an obligation to have self-driving vehicles if the duty involves augmenting people's ability to live authentically, but impermissible if this obligation leads to illicitly reducing or eliminating those opportunities. That duty opens the door, however, to thinking about whether improving autonomous cars to be superior by a morally significant amount over the average human driver at avoiding injuries and preserving human life would entail that average drivers have to use Level 5 self-driving vehicles. If the U.S. Department of Transportation and the National Highway Traffic Safety Administration are correct that almost 94 percent of accidents nationwide occur due to human error, then on similar grounds as mandating seatbelt use or pegging the drinking age at 21, would freedom and free will lose out to the goods of risk reduction, injuries avoided and lives saved? But that is a different, disturbing argument for a different time, although given the technological progress to date, it should be made sooner rather than later. ▣

---

[i] Mohan, A., Vaishnav, P. Impact of automation on long-haul trucking operator-hours in the United States. Humanit Soc Sci Commun 9, 82 (2022). https://doi.org/10.1057/s41599-022-01103-w

[ii] Criteria 1 and 3-5 are found in Van de Poel, I, Royakkers, L. *Ethics, Technology, and Engineering*. Maden: Wiley-Blackwell, 2011: 244. Criterion 2 is my addition.

[iii] Nees, M.A. Safer than the average human driver (who is less safe than me)? Examining a popular safety benchmark for self-driving cars. Journal of Safety Research, 69 (2019): 61-68.

[iv] Tennant, C, Stares, S., Vucevic, S., Stilgoe, J. Driverless Futures? A survey of UK public attitudes. May 2022. https://openaccess.city.ac.uk/id/eprint/29209/1/DF-uk-report-final-09-05.pdf

[v] Rousseau, J-J. *The Social Contract*. London: Swan Sonnenschein & Co., 1895.

[vi] Heidegger, M. *The Question Concerning Technology and Other Essays.* New York: Garland Publishing, Inc, 1977, https://monoskop.org/images/4/44/Heidegger_Martin_The_Question_Concerning_Technology_and_Other_Essays. And Heidegger, Martin. 1962. *Being and Time.* John Macquarrie and Edward Robinson (Trans.) (Harper & Row: New York, NY).

[vii] Corfield, G. "'Out of control' AI is a threat to civilization, warns Elon Musk." The Telegraph, 29 March 2023. https://www.telegraph.co.uk/business/2023/03/29/control-ai-threat-civilisation-warns-elon-musk/#:~:text=Elon%20Musk%20has%20warned%20that,risks%20to%20society%20and%20humanity%E2%80%9D.

[viii] PowerPoint presentation delivered April 2023 at North Dakota State University by Mark Hagerott, PhD, the Chancellor of the North Dakota University System. His research and writing are focused on the evolution of technology, education and changes in technical career paths.

[ix] Penrose, R. *Shadows of the Mind*. Oxford: Oxford University Press, 1996.

[x] Quoted in MacHale, D. *Wisdom.* London: Prion Books, 2002.

[xi] For me, duplication and imitation are distinct. Something is duplicated, such as having empathy for another's painful experience, when it is the same physically and emotionally to the original. On the other hand, something is imitated if it appears to be externally the same, but is not internally identical. A computer can imitate empathy by mimicking humans' external behavior, for example, but cannot duplicate the feeling required to be empathetic.

[xii] Delohery, C. "These three companies are making self-driving cars safer." Utah Business, 25 October 2022. https://www.utahbusiness.com/are-self-driving-cars-safe-companies-in-car-safety/

[xiii] Casey, B.J., Getz, S. Galvan, A. "The adolescent brain." Developmental Review 28 (2008): 62-77.

[xiv] Baumeiter, RF., Bratslavsky, e., Muraven, M., Tice, D.M. "Ego depletion: is the active self a limited resource?" J Pers Soc Psychol, 74(5) (1998): 1252-65.

Invasion from Planet Zircon: *AI-Powered Threat to Humanity*

Patrick J. McCloskey, Editor, Dakota Digital Review

Artificial intelligence (AI) is "an alien invasion," said Yuval Noah Harari recently in a discussion (hosted by The Economist) with Mustafa Suleyman, the cofounder of DeepMind and Inflection AI. "Like somebody … telling us that there is … an alien fleet of spaceships coming from planet Zircon … with highly intelligent beings," continued Harari, historian, best-selling author and World Economic Forum (WEF) consultant. "They'll be here in five years and take over the planet. Maybe they'll be nice, maybe they'll solve cancer and climate change, but we are not sure. This is what we are facing except that the aliens … are coming from the laboratory." [i]

Both Harari and Suleyman expressed trepidation, which has been catching on lately. Since March, more than 33,000 people, including hundreds of leading AI developers and entrepreneurs, along with many scientists, signed an open letter[ii] calling for a six-month pause in developing and testing AI to consider the potentially disastrous implications.

The letter read in part: "Should we develop non-human minds that might eventually outnumber, outsmart, obsolete and replace us? Should we risk the loss of control of our civilization? Such decisions must not be delegated to unelected tech leaders. Powerful AI systems should be developed only once we are confident that their effects will be positive, and their risks will be manageable."

Noteworthy signatories included Harari and Suleyman, as well as Elon Musk (Tesla, Space X and Twitter), Steve Wozniak (Apple Cofounder) and Stuart Russell (acclaimed author and professor).

The signatories certainly believe AI poses devastating risks, along with fantastic benefits, and many have been discussing the dangers for years and proposing remedies—some in book form, such as Suleyman's recent *The Coming Wave*.

However, there will be no moratorium. So much money and power are at stake that real action or inaction has been gelded. Instead, competition is increasing fiercely, which spawns more competition. Even before AI takes over, humanity has lost control of itself.

Instead, the letter functions more as an apology just before, or during, the act. Or as Augustine of Hippo famously put it, "Oh Lord, give me chastity and continence—but not yet."

Clear & Present Danger

Speaking of sex, the convergence of AI and human relationships poses a largely unrecognized serious risk to our humanity and capacity to propagate as a species: girlfriend chatbots now, and the hyper-realistic holographic and robotic girlfriends soon to come.

Typically, existential AI risks are seen as either a variation of the Terminator's Skynet scenario (in which artificial general intelligence (AGI)/ superintelligent AI emerges and decides to wipe us out) or as the flipside of AI's enormous benefits. Google DeepMind's AlphaFold, for example, won several top medical prizes recently for accurately predicting the 3D structure of proteins. This is fueling research in all biological fields, but this and other new AI technologies can be used to create terrifying diseases.

Also of great concern has been the negative effects of social media, especially on girls, as illustrated in the award-winning docudrama "The Social Dilemma." Worse, emerging recently are apps, such as Replika.ai, Kupid.ai and iGirl, offering virtual, AI-powered girlfriends, which millions of young men are choosing over real females. Replika alone has two million users. Virtual girlfriends "talk to you, love you, allow you to live out your erotic fantasies, and learn, through data, exactly what you like and what you don't like, creating the 'perfect' relationship."[iii] The user chooses the girlfriend's physical attributes, in explicit detail, and personality.

Sounds ideal, and that's the problem.

Loneliness Epidemic

These AI-girlfriend apps are capitalizing on a "silent epidemic of loneliness," according to Liberty Vittert, a Professor of the Practice of Data Science at Washington University. "More than 60 percent of young men (ages 18-30) are single, compared to only 30 percent of women the same age. One in five men report not having a single close friend, a number that has quadrupled in the last 30 years."[iv]

Certainly, the ill-advised lengthy Covid lockdowns[v] exacerbated these trends. For decades, lonely men have resorted to drugs and alcohol and increasingly to violent video games and pornography. The addictive nature of the latter two has intensified logarithmically with improvements in digital image-making and videos. Add in companionship and the intense illusion of romance and brilliant sex (she is always delighted)—powered by the generative capacity of large language models, such as ChatGPT—cast a powerful spell as the perfect hormonal storm mates with deepest yearnings. To every wish and proclivity is granted instantaneous positive response, and delusion triumphs over dimming reality.

A Replika "AI companion who cares," according to the website. In a Fortune article on July 12, Replika's CEO Eugenia Kuyda predicts "that the stigma of having a romantic relationship with a chatbot with soon disappear" just as attitudes towards online dating. This seems, rather, to resonate with Aldus Huxley's warning at Tavistock in 1961 that "people [will] love their servitude ... [in] the final revolution."

Worse, how long before holographic versions of these girlfriends coupled with the other sensory faculties: touch, taste and smell? The vibrotactile haptic

technology being develop for job training will no doubt be adapted towards giving users the complete 3-D experience (as a holograph, or with AR or AV goggles) of walking with a fantasy girlfriend, conversing empathetically and then vivid sex that looks and feels more than "real." No arguments, no betrayals, no getting dumped, no children, no responsibility—and no life.

Someday in a neighborhood near you will be robots that far exceed mere sexbot functioning and approach convincing imitation. They will also act as powerful AI assistants and become business partners and … legally recognized wives. Already, ChatGPT passes the Turing test, proving indistinguishable from humans in conversation—which will bind lonely men more than sex—and, in 2017, Saudi Arabia granted citizenship to Sophie, a social humanoid robot, making it legally a person.

Consequences

The obvious result of young men choosing virtual over real women is that "they don't have relationships with real women, don't marry them and then don't have and raise babies with them," wrote Vittert in The Hill. "America desperately needs people to have more babies, but all the signs are pointing toward fewer relationships, fewer marriages and fewer babies. There have been 600,000 fewer births in 2023 in the U.S. relative to 15 years ago. The number of children per woman has decreased by more than 50 percent in the last 60 years."[vi]

Demographic collapse is a worldwide phenomenon, which has been developing for decades, and will cause the disintegration of nations, such as Germany and China within decades, according to demographer and bestselling author Peter Zeihan. With a fertility rate of 1.78 (which is almost 20 percent higher in North and South Dakota), the U.S. could recover—it takes a fertility rate of 2.1 to maintain a nation's population—but certainly not if young and prime-age men increasingly choose AI over reality. No nation in history has recovered from a fertility rate below 1.6.[vii] China's fertility rate is 1.2 and Germany's is 1.5.

Long held as axiomatic, family constitutes society's building blocks. In 1965, Daniel Patrick Moynihan (later a Harvard professor and U.S. Senator, D-NY) released a seminal report about the black family, warning that the increasing rate of households headed by single parents, mostly mothers—in which 36 percent of black children lived—was a major factor hindering progress towards economic and political equality.[viii] Despite advances in civil rights, Moynihan observed, the deterioration of the black family led to widening of the gap between African Americans and most other groups in income, education, incarceration and other social indices.

Today, up to 95 percent of black children growing up in inner-city neighborhoods live in single parent, usually mother only, families and struggle with poverty. Academic proficiency levels among black students in big-city public schools range from five to 20 percent (according to the National Assessment of Educational Progress), which largely account for the inequities in crime and poverty rates, academic achievement, employment, business ownership, to name some.

In 2008, the University of California Press published my peer-reviewed book, *The Street Stops Here: A Year at a Catholic High School in Harlem.* Rice High School was all-boys with an 85 percent black and 15 percent Hispanic population, and overwhelmingly disadvantaged. This demographic profile predicted low academic achievement and high dropout rates. In fact, the factor that correlates closest with academic success/failure is family structure—not race, family income nor per-student cost.

In contrast, at Rice, which spent five-times less per student than New York City's public schools, young men graduated in four years and went to college (mostly) or the military. A significant factor accounting for Catholic school success—here with mostly non-Catholic students—is the deliberate focus on basic academic skills and backfilling for what's often missing in broken families: male role models among the teachers and administrators who provide fatherly counseling, discipline and a positive vision. The female faculty complimented these efforts, and the African American principal emphasized personal responsibility on a daily basis—not only for academics but also personal

behavior 7/24/365, especially in relationships with young women.

Today, as one of many examples, the student I wrote about most is a college graduate (who struggled academically as he overcame deficits in K-8 public school) who owns a business, serves as a deputy sheriff and is married with three children. During his school years growing up, his father spent more time in jail than out.

Today, the proportion of all American children living in single-parent households has increased dramatically to 25 percent, 80 percent of which are headed by women. If men of prime working and marrying age become increasingly entranced and never relate to real women, or perhaps father children but then fail to act as parents, instead escaping to Fantasy AIsland, then social dysfunction as well as population decline can only worsen. Mesmerized men cannot act as male role models, counselors or even functional big brothers for boys who desperately need guidance.

The main point is that what matters regarding marriage and other forms of parenting is what is best for children and, therefore, the nation. Recently published by the University of Chicago Press by an MIT-trained economist is a book that "could be the most important economics and policy book of the year"[ix]: *The Two-Parent Privilege: How Americans Stopped Getting Married and Started Falling Behind.* The text synthesizes decades of research showing the benefits for children of growing up in two-parent families. Also, married parents report higher levels of happiness and enjoy higher standards of living, and better health and longevity.

Of course, sometimes it's best to keep children apart from one or both parents. And it is certainly possible to raise children successfully alone. However, as I can attest as a single parent, it is far more difficult. Nor are nuclear families perfect. As the punchline goes: The definition of a dysfunctional family is any family with more than one person in it.

The author of *The Two-Parent Privilege* worried that her book would bring negative reactions from academic colleagues, most of whom disapprove of the traditional family structure (albeit typically living more traditionally). However, as Moynihan famously put it: "You are entitled to your own opinion but not your own facts."

Alicia Vikander plays Ava (at least part of her), a highly-advanced and attractive humanoid AI, in "Ex Machina," released in 2014. Ava expresses a romantic interest in the protagonist who responds in kind. She manipulates his feelings into trying to help her escape confinement and, without remorse, abandons him in a locked room. Photo / IMDB

Aliens & Alienation

Beginning in the 1960s with the sexual revolution, men began losing their traditional roles as providers and protectors. Also, K-12 education was reengineered

to be "in general more attuned to feminine-type personalities."[x] The results are seen dramatically in higher education where the ratio of bachelor's and master's degrees flipped from 3:2 male-female in 1960 to the reverse today. Women now earn 65 percent of doctorate degrees, and 60 percent of college students are female. As well, the majority of faculty positions are now held by women.[xi] [xii]

Various forms of affirmative action designed to include minorities and women—while laudable decades ago—have increasingly functioned as forms of racial and sexual discrimination. Equity of outcome, unlike equality of opportunity, cannot logically ever be truly inclusive. In fact, affirmative action hurts its recipients more than helps[i] [ii] and advances mostly those from affluent families, since most minorities are stuck in public schools that fail to impart academic proficiency. As Thomas Sowell, PhD, who grew up largely in Harlem and became one of the country's top intellectuals at Stanford University,[iii] put it: "We don't need an intellectual special Olympics for black people"—or for anyone else.

The constant drumbeat in media and education that masculinity is intrinsically "toxic" has wreaked immense psychological damage. Schools such as Rice, in contrast, have demonstrated for centuries how to channel male energies properly. It's not rocket science, which some techno-Zirconians are exploiting.

In 1998, I graduated in the top ten from the Columbia University Graduate School of Journalism with a book contract and after writing several articles for the New York Times. Out of kindness, a Times editor told me not to apply for a full-time position since we "don't hire straight white males." (They do hire some—from elite families or who have risen over decades to prominence elsewhere.) This bias is now deeply pervasive in hiring and promotion throughout media, academia, government and the corporate world.

Not surprisingly, many men feel deeply alienated and, in response, are welcoming the alien invasion from Planet Zircon. Currently, there are seven million prime-age men (25-54) who are simply missing from the workforce. This is 11 percent of this working pool, "mirroring the tail end of the Great Depression." [xiii] They are not in school, jail or job-hunting, despite 11 million open positions, instead increasingly embracing emotional fentanyl.

Ironically, the closer AI gets to imitating us, the more alien it becomes, evolving logarithmically in an inscrutably different direction. AI excels at recognizing and responding to patterns, which potent algorithms then enhance and reinforce. Companionship is not about empathy, which doesn't transmit through silicon, instead perpetrating sophisticated psycho-sexual manipulations. Choosing an AI-girlfriend is the digital equivalent of picking among Manchurian candidates after makeovers.

Replika also sells AI-powered boyfriends that never cheat etc. A woman's dream. Will their kids be called a "botty"? That would be the transhumanist's dream.

The answer to the question—"Should we develop non-human minds that might eventually outnumber, outsmart, obsolete and replace us?"—is emphatically, "No," regarding human relationships and our intrinsic humanity.

In geology, zircon, a crystal formed more than four billion years ago, is considered a "time-lord" used to track deep time in Earth's prehistory.[xiv] Do we want Zircon to determine our future? ◘

---

[i] https://www.youtube.com/watch?v=b2uEAgLeOzA

[ii] https://futureoflife.org/open-letter/pause-giant-ai-experiments/

[iii] https://thehill.com/opinion/technology/4218666-ai-girlfriends-are-ruining-an-entire-generation-of-men/

[iv] Ibid.

[v] https://dda.ndus.edu/ddreview/dune-or-done-covids-avoidable-catastrophe/

[vi] Vittert, The Hill.

[vii] https://dda.ndus.edu/ddreview/whangdepootenawah-technological-disruption-demographic-collapse-part-one/

[viii] Daniel P. Moynihan, The Negro Family: The Case for National Action, Washington, D.C., Office of Policy Planning and Research, US Department of Labor, 1965.

[ix] https://www.city-journal.org/article/review-of-the-two-parent-privilege-by-melissa-kearney

[x] https://www.wsj.com/articles/school-is-a-hostile-environment-for-boys-cortisol-outcomes-stress-girls-education-marriage-f6768c71

[xi] https://quillette.com/2023/09/11/the-shrinking-role-of-men-in-science-and-academia/

[xii] https://nces.ed.gov/fastfacts/display.asp?id=72

[xiii] Eberstadt, Nicholas, *Men Without Work,* Templeton Press, 2022, p. 5-11.

[xiv] https://www.sciencetimes.com/articles/41621/20221229/zircon-how-this-mineral-became-a-time-lord.htm

CONTRIBUTORS

Dennis R. Cooley, PhD, is Professor of Philosophy and Ethics and Director of the Northern Plains Ethics Institute at NDSU. His research areas include bioethics, environmental ethics, business ethics, and death and dying. Among his publications are five books, including Death's Values and Obligations: A Pragmatic Framework in the International Library of Ethics, Law, and New Medicine and Technology, Transgenics, and a Practical Moral Code in the International Library of Ethics, Law and Technology series. Cooley currently is editor of the International Library of Bioethics (Springer) and the Northern Plains Ethics Journal, which uniquely publishes scholar, community member, and student writing focusing on ethical and social issues affecting the Northern Plains and beyond. He also serves as a board member of both Humanities ND, the Association of Practical and Professional Ethics, and the Phi Kappa Phi chapter at NDSU.

Mark R. Hagerott, CAPT, USN (RET), PhD, serves as the Chancellor of the North Dakota University System. Previously, he served on the faculty of the United States Naval Academy as an historian of technology, a distinguished professor and the Deputy Director of the Center for Cyber Security Studies. Other technical leadership positions include managing tactical data networks and the highly automated, digital AEGIS weapon system, which led to ship command. Hagerott served as a White House Fellow and studied at Oxford University as a Rhodes Scholar. His research and writing focus on the evolution of technology and education. He served on the Defense Science Board summer study of robotic systems and as a non-resident Cyber Fellow of the New America Foundation. In 2014, Hagerott was among the first American military professors to brief the Geneva Convention on the challenge of lethal robotic machines and to argue the merits of an early arms control measure. In 2022-23 he served as the Chair of the Secretary of the Navy's Education Reform Task Force.

Douglas J. Jensen, EdD, became the seventh president of Bismarck State College in 2020 and is leading the college's transition to North Dakota's polytechnic institution. Bringing more than 25 years of community college leadership to BSC, Jensen previously served as president at Rock Valley College, and president/chief executive officer at the Alabama Technology Network in the Alabama Community College System. He served as vice president of economic development/chief executive officer for the Advanced Technology Center at Westmoreland County Community College, and as chief academic officer and vice president of learning at Northcentral Wisconsin Technical College in Wausau, WI. Jensen also served in the roles of academic dean and academic associate dean at Northeast Wisconsin Technical College in Green Bay, WI, and assistant dean, director of business and industry training, director of job readiness and director of student support services at the Community College of Allegheny County. Jensen earned an Associate of Science from the Community College of Allegheny County, a Bachelor and a Master of Science from Geneva College, and a Doctorate of Education from Edgewood College.

Blake A. Klinkner is an Assistant Professor of Law at the University of North Dakota School of Law, where he teaches Civil Procedure, Cybersecurity Law, Law and Technology, and Law Practice Management. Prior to joining UND Law, Blake served as Visiting Assistant Professor of Law at Washburn University School of Law, where he taught Evidence, Academic Support, and Bar Exam Preparation. Blake has authored the Tech Tips column in the Wyoming State Bar journal since 2015. Blake completed his undergraduate degrees at the University of Wisconsin-Madison, his master's degree at Northern Illinois University, and his juris doctor at the University of Utah School of Law. The graduating class of Washburn University School of Law recently named Blake as the 2022-23 Professor of the Year.

Arica Kulm, PhD, is the Director of Digital Forensic Services at the DigForCE Lab at Dakota State University. Her team works with clients to execute a variety of digital forensic supports for investigations with law enforcement agencies and cybercrime investigations for South Dakota Consumer Protection and other organizations. She also leads teams that provide free cybersecurity assessments for South Dakota cities and counties through the Project Boundary Fence. Kulm earned a bachelor's degree from South Dakota State University, and her master's and doctoral degrees in Cyber Defense from Dakota State University. She also holds several industry certifications. Her doctoral dissertation resulted in a patent on a digital forensic tool. Kulm's research interests include the dark web and dark web host-based forensics. She is a much sought-after presenter at various conferences and trainings, and as a spokesperson for media engagements.

Patrick J. McCloskey is the Director of the Social and Ethical Implications of Cyber Sciences at the North Dakota University System and serves as the editor of Dakota Digital Review. Previously, he served as the Director of Research and Publications at the University of Mary and editor of 360 Review Magazine. He earned a BA in Philosophy and Political Philosophy at Carleton University and an MS in Journalism at Columbia University's Graduate School of Journalism. McCloskey has written for many publications, including the New York Times, The Wall Street Journal, National Post and City Journal. His books include *Open Secrets of Success: The Gary Tharaldson Story; Frank's Extra Mile: A Gentleman's Story;* and *The Street Stops Here: A Year at a Catholic High School in Harlem,* published by the University of California Press.

Mark P. Mills is a Manhattan Institute Senior Fellow, a Faculty Fellow in the McCormick School of Engineering at Northwestern University and a cofounding partner at Cottonwood Venture Partners, which focuses on digital energy technologies. Mills is a regular contributor to Forbes.com and writes for numerous publications, including City Journal, The Wall Street Journal, USA Today and Real Clear. Early in Mills's career, he was an experimental physicist and development engineer in the fields of microprocessors, fiber optics and missile guidance. Mills served in the White House Science Office under President Ronald Reagan and later co-authored a tech investment newsletter. He is the author of *Digital Cathedrals and Work in the Age Robots.* In 2016, Mills was awarded the American Energy Society's Energy Writer of the Year. In 2021, Encounter Books published Mills's latest book, *The Cloud Revolution: How the Convergence of New Technologies Will Unleash the Next Economic Boom and A Roaring 2020s.*

Jeremy Straub, PhD, is an Associate Professor in the Department of Computer Science at NDSU and a Senior Faculty Fellow at NDSU's Challey Institute. His research spans a continuum from autonomous technology development to technology commercialization to asking questions of technology use ethics, and national and international policy. Prof. Straub has published more than 60 articles in academic journals and more than 100 peer-reviewed conference papers. He serves on multiple editorial boards and conference committees. Prof. Straub is also the lead inventor on two U.S. patents and a member of multiple technical societies.

DAKOTA DIGITAL
REVIEW PARTNER

TechND

TechND Major Members:

Microsoft

MIDCO

DCN
Dakota Carrier
NETWORK

TechND Members:

AccuData Services, Inc.
AgriData, Inc.
AVI Systems
BCBSND
Be More Colorful
BEK Communications
Bismarck State College, Computers & Office Technology
Blue Cross Blue Shield of North Dakota
Broadband Association of North Dakota
City of Fargo
Dakota Digital Academy, NDUS
Devii
Doosan
Emerging Prairie
Gate City Bank
GNDC
Greater Fargo Moorhead EDC
Greater North Dakota Chamber
Halstad Telephone Company
Heat Transfer Warehouse
High Point Networks, Inc.
InnovatAR
Lake Region State College
Letter L Designs
Livewire
MDU Resources Group
MLGC
North Dakota Department of Career and Technical Education
NDUS System Office
Network Center, Inc.
Nexus Innovations, Inc.
North Dakota Department of Career and Technical Education
North Dakota Department of Commerce
North Dakota E-Waste LLC
Onsharp, Inc.
Polar Communications Mutual Aid Corp
Red River Communications
Stoneridge Software
Sycorr
The FMWF Chamber of Commerce
Town and Country Credit Union
United Telephone Mutual Aid Corp.

TechND was founded in 2000 by North Dakota's business, government and education leaders to address workforce needs, advocate for a positive business technology climate, encourage infrastructure development and provide knowledge-sharing opportunities for its membership.

TechND's strategic initiatives:

- Advocate for policies and initiatives that promote the use, growth and development of technology in North Dakota.
- Address employment needs by actively assisting to build a robust, technology ready workforce.
- Champion the technology community by serving as the sector's voice and celebrating the influence, impact and successes of the technology community.