INSTITUTE FOR DEFENSE ANALYSES

Megatrend Issues in Artificial Intelligence and Autonomous Systems

Clifford G. Lau

Brian A. Haugh

August 2018

IDA Non-Standard
NS D-9144

INSTITUTE FOR DEFENSE ANALYSES
4850 Mark Center Drive
Alexandria, Virginia 22311-1882

The Institute for Defense Analyses is a non-profit corporation that operates three federally funded research and development centers to provide objective analyses of national security issues, particularly those requiring scientific and technical expertise, and conduct related research on other national challenges.

About This Publication

This work was conducted by the Institute for Defense Analyses (IDA) under Central Research Project C5190, "Artificially Intelligent Autonomous Systems." The views, opinions, and findings should not be construed as representing the official position of either the Department of Defense or the sponsoring organization.

Acknowledgements:
Robert M. Rolfe, David A. Sparrow

For more information:

Clifford G. Lau, Project Leader
clau@ida.org, 703-933-6525

Margaret E. Myers, Director, Information Technology and Systems Division
mmyers@ida.org, 703-578-2782

Megatrend Issues in Artificial Intelligence and Autonomous Systems

Clifford G. Lau (clau@ida.org)
Brian A. Haugh (bhaugh@ida.org)

August 2018

INSTITUTE FOR DEFENSE ANALYSES

Megatrends are sustained developments that fundamentally impact business, economy, society, cultures, and personal lives. Recent advances in artificial intelligence (AI) will enable autonomous systems (AS), with far-reaching implications in both the civilian sector and defense. AI-enabled robots will perform difficult and dangerous tasks that require human-like intelligence. Self-driving cars will revolutionize automobile transportation systems and reduce traffic fatalities. Big-data analytics using AI techniques will make human-like decisions to improve governmental social services, health care, criminal justice, and the environment. AI-enabled autonomous robotic soldiers, aerial drones, and underwater and land vehicles will perform military missions. These revolutionary technological advances will have significant impacts on the economy, military, and society. We are seeing a whole new generation of AI and AS that will change, in unforeseen ways, how we live, work, play, and fight wars. However, for the public and military to adopt AI and AS, society and military must have confidence that these systems are trustworthy and safe. A number of important issues are awaiting policymakers, including research and development, workforce development, safety, cybersecurity, ethics, regulations, and automated warfare.

1. Introduction

Megatrends are global and sustained movements and developments that fundamentally change business, economy, society, cultures, and personal lives. In science and technology, megatrends often are the results from the development of revolutionary technologies, such as the Internet. Artificial intelligence (AI) technologies and their application to autonomous systems (AS) have emerged as one such megatrend that is expected to have wide-ranging influence on human society going forward. AI and AS have been very much in the news lately. Recent advances in AI applications, such as self-driving cars, smart personal assistants, image/video understanding, and game playing, have captured the public's imagination and the interests of governments, industries, and militaries across the world. Although these advances justify a widespread enthusiasm for AI and AS, there is need for caution in preparing for their anticipated and unanticipated impacts on the way we live, work, play, and fight wars in the future. Looking back at the history of AI and AS, as well as looking forward to their expected and potential future applications and consequences, will lead to a better understanding of their potential impacts and how to prepare for them. After providing this context, we review critical issues that need to be addressed by legislators and policymakers for successful development and deployment of these technologies, including the conduct of research and development, workforce development, cybersecurity, ethics and regulations, and automated warfare.

1.1. Artificial Intelligence

The term *artificial intelligence* was first coined by John McCarthy at the Dartmouth Conference on the topic in 1956. The term was not formally defined at that time, although the proposal for the conference described its intent to broadly address the machine simulation of intelligence:

> The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it. An attempt will be made to find how to make machines use language, form abstractions and concepts, solve kinds of problems now reserved for humans, and improve themselves.[1]

Subsequently, different authors have proposed varying definitions, upon which there has never been widespread agreement, although at the core they are all about the capability of artificial systems (machines) to exhibit "intelligent" behavior. However, the bar for what counts as intelligent behavior has been raised over the years as some have come to regard the achievements of computers in natural language processing, image recognition, game playing, reasoning, planning, diagnosis, and AS to fall short of genuine "intelligence." Although it is widely recognized that today's systems using AI technologies are very far from approaching human-level general intelligence, many of them are reasonably considered to exhibit some narrow types of intelligent behavior.

The field of AI was introduced during the early days of digital computers. In the beginning, AI researchers wrote computer programs that were, to ordinary people, simply astonishing and were recognized as showing some form of intelligence. Soon the computers were winning at checkers, solving word problems in algebra, proving logical theorems, and speaking English. Research in the United States was heavily funded by the Department of Defense (DoD). At the time, the outlook for AI was promising, and it was expected to be able to soon solve problems with human-like intelligence. However, the initial excitement became disappointment as researchers recognized the limitations of then current AI technologies when faced with tasks that required human-level intelligence.

Over the more than six decades since the initial Dartmouth AI conference, AI has experienced several hype cycles, with interest waxing and waning. At times, the technology delivered great promise, and at others, it failed to meet expectations. In the early 1980s, AI research was revived by the commercial success of so-called expert systems, which were a form of AI program that simulated the knowledge and analytical skills of human experts. By the mid-1980s, the market for AI had reached over a billion dollars and the U.S.

---

[1] McCarthy, J., M. Minsky, N. Rochester, and C.E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", August 31, 1955. Reprinted in AI Magazine, Winter 2006, 12. https://www.aaai.org/ojs/index.php/aimagazine/article/download/1904/1802

Government had restored some funding for AI research. However, the difficulty for expert systems to solve real-world problems that require human-level intelligence again surfaced, as expert systems were recognized to be brittle – providing bad results when applied to problems outside their narrowly circumscribed applications. This, together with the collapse of the market for specialized AI computers such as the LISP Machine,[2] led to an AI "winter", in which researchers avoided the use of the term "AI" in their proposals if they wanted to get funded. AI had, once again, fallen into disrepute, and a second, longer-lasting hiatus began. Still, AI research continued at a low level by some dedicated computer science researchers.

By the early 2000s, the persistent research began to show some promising results, and AI began to be used for logistics, data mining, medical diagnosis, and other applications. Enabled by increasingly powerful digital computers due to Moore's Law, researchers developed more powerful AI algorithms. A breakthrough point occurred in 1997 when IBM's Deep Blue computer chess program defeated the reigning world chess champion Garry Kasparov. Recent successful AI demonstrations – such as IBM's Watson competing against a human Jeopardy champion and winning; Google besting a human champion in the ancient Chinese board game, Go; and the popular demonstrations of self-driving cars – have all contributed to renewed and increased focus on and excitement about AI.

1.2. Autonomous Systems

AS are machines that operate without the active intervention of a human operators. The technologies used in AS often include sensors, computers, and AI. Sometimes used interchangeably with autonomous robots, AS encompass a large number of computerized machines such as unmanned vehicles on land, at sea, in the air, and in space. Similar to AI, development in AS began in the early days of digital computers. In the beginning, simple robots were programmed to perform tasks autonomously within the confines of the factory environment. Soon, factory robots were deployed in assembly lines, such as the familiar robotic welding arms in automobile factories. Gradually, developers were able to create mobile robots that could freely navigate their way in the environment, such as the iRobot and Electrolux vacuuming robots in 2002. DoD accelerated the development of unmanned aerial vehicles (UAVs), or pilotless drones, for use in surveillance and reconnaissance. The UAVs were capable of flying their entire mission without any human interaction, except possibly for the take-off and landing where a person intervened using radio remote control. There are various levels of autonomy, as discussed in NIST Special Publication 1011-I-2.0, *Autonomy Levels for Unmanned Systems (ALFUS) Framework*. The Society of Automotive Engineers (SAE) International broadly defines five levels of autonomy for self-driving cars: 1–operator assistance, 2–partial automation, 3– conditional automation,

[2] LISP machines were special purpose machines developed to process symbolic AI programs using the LISP programming language. They were subsequently superseded by faster general purpose CPUs.

4–high automation, and 5–full automation. Vehicles sold today mostly correspond to levels 1 and 2, while fully autonomous ground vehicles (i.e., self-driving cars) are not expected to be widely deployed within five years, although that could happen sooner given the rapid pace of development.

Fully autonomous vehicles have been an international pursuit for many years since 1977. In 2004, the U.S. Congress authorized the Defense Advanced Research Projects Agency (DARPA) to offer prize money for the first Grand Challenge to facilitate robotic development, with the ultimate goal of making one-third of ground military forces autonomous by 2015. The first competition of the DARPA Grand Challenge, held on March 13, 2004, in the Mojave Desert region, required a driverless car to travel a 150-mile route. Unfortunately, none of the robot vehicles finished the route. In the second competition in 2005, five autonomous vehicles completed the course, with the first place prize going to Stanford Racing Team. The third competition, known as Urban Challenge, took place on November 3, 2007. The course involved 60 miles of urban roads to be completed in less than 6 hours. Six teams successfully finished the entire course, with Carnegie Mellon University's Tartan Racing Team claiming the $2 million prize. Turning to humanoid robotics, the DARPA Robotics Challenge was competed in October 2012. The goal of the competition was to develop ground robots capable of executing complex tasks in dangerous environments. The Robotics Challenge was followed by the FANG (Fast Adaptable Next Generation Grand Vehicle Mobility/Drivetrain) Challenge in 2013. These autonomous vehicle and robotic challenges laid the foundation for much of the development of autonomous robotic systems, although they did not achieve the optimistic goals that Congress envisioned in 2004.

Recent advances in AI such as machine learning and computer vision have contributed to greater intelligence in AS. Advances in sensor technology such as lidars, radars, and advanced imaging sensors have also contributed to the development of AS. Modern self-driving cars generally use Bayesian simultaneous localization and mapping (SLAM) algorithms, which fuse data from multiple sensors and an offline map into current location estimates and map updates. The algorithm for detection and tracking of other moving objects (DATMO), which handles things such as cars and pedestrians, is a variant being developed at Google. Simpler systems may use roadside real-time locating system (RTLS) beacon systems to aid localization. Typical sensors include lidar, stereo vision, GPS, and inertial measuring units (IMU). Visual object recognition uses machine vision, including convolutional neural networks (CNN) and deep learning (DL).

Recent demonstrations of self-driving cars by Google, Tesla, and others have caused much excitement. The potential benefits of autonomous cars include reductions in automobile fatalities, energy consumption, and operational costs. Self-driving cars are envisioned to result in significant reduction in traffic collisions and injuries, and thus in less need for high-cost insurance. Autonomous cars are predicted to increase traffic flow,

provide mobility to the elderly and disabled, lower fuel consumption, and reduce the need for parking spaces.

With the advances in AI and AS, we are seeing a new era of advanced technology. There are obstacles to wide adoption of the AI and AS technology, including consumer safety concerns, disputes concerning liability, implementation of a workable legal framework, establishment of government regulations, risk of loss of privacy, cybersecurity concerns, and loss of jobs due to automation. To oversee the legislative agenda, the U.S. House of Representative has formed the Robotics Caucus and the Artificial Intelligence Caucus. The Robotics Caucus has held workshops and hearings on the possible displacement of workers and loss of jobs due to increased automation. The AI Caucus has also held workshops and hearings on the potential benefits and pitfalls of AI and on educating the public on AI facts and fictions.

Potential deployment of self-driving cars in public streets has prompted Congress to provide new regulatory tools to the National Highway Traffic Safety Administration (NHTSA) to oversee autonomous vehicles. The federal government shares motor vehicle regulations with state governments, with the federal government responsible for vehicle safety and state governments responsible for licensing and registration. Although NHTSA has statutory authority to regulate all types of motor vehicles, traditional standard-setting processes and regulations would take too many years to put in place to respond to the rapid advances in self-driving cars. Nearly 25 states have enacted laws on different aspects of autonomous vehicle deployment. On September 6, 2017, the House of Representatives passed H.R. 3388, SELF DRIVE Act, which preempted states from regulating the design of autonomous vehicles and expanded NHTSA's authority to grant exemptions from its conventional standards. The legislation required each autonomous vehicle manufacturer to submit safety assessment certification showing how safety is addressed and to develop and publicize to consumers their cybersecurity and data privacy plans.

The rapidly developing AI and AS technology has broad implications for society, culture, the economy, and the military. In a summer study[3] on autonomy, the Defense Science Board (DSB) made recommendations to improve the future adoption and use of AS in the military. The study recommended a set of experiments and prototypes that would demonstrate clear operational value to the military and would serve as pilot projects to help refine and institutionalize the recommendations. Since the rapid developments in AI and AS are taking place in the commercial sector, DSB recommends that DoD take steps to engage the non-defense research and development community to speed DoD's access to

[3] Office of the Under Secretary of Defense for Acquisition, Technology and Logistics, *Report of the Defense Science Board Summer Study on Autonomy*, June 2016. https://www.acq.osd.mil/dsb/reports/2010s/DSBSS15.pdf

the emerging AI and AS research results and to identify areas in which DoD investment is needed to fully address DoD missions.

2. Research and Development

Greater federal investment in AI and AS research and development (R&D) is essential to stimulate the economy, maintain U.S. competitiveness, create high-value jobs, and improve government services. The recent acceleration of successful AI applications makes it timely to focus R&D investment in more capable AI systems and also to maximize societal benefits and mitigate any associated risks. R&D is needed to overcome possible obstacles to deployment of AS such as self-driving cars.

Presently, the NITRD (Networking, Information Technology, Research and Development) National Coordination Office (NCO) coordinates federally funded R&D in a number of IT program areas, including Intelligent Robotics and Autonomous Systems (IRAS).[4] IRAS R&D focuses on advancing intelligent robotics and AS that complement, augment, enhance, or emulate human physical capabilities or embodied human intelligence. Examples include robotics hardware and software design, application, and practical use; machine perception; intelligent cognition, adaptation, and learning; mobility and manipulation; human-machine interaction; distributed and networked robotics; increasingly autonomous systems; and related applications. FY16 IRAS funding was $225 million, or 5% of the NITRD budget. DoD's contribution was $101.8 million, a significant 48% of the IRAS investment. The planned IRAS investment in FY17 was $220.5 million, a decrease of $4.5 million, with DoD contributing a stable $102.9 million. Given the potentially significant societal benefits from AI/AS R&D investment, the federal government should greatly increase – not decrease – IRAS R&D funding.

AI is a transformative technology that holds promise for tremendous societal and economic benefits. The National AI R&D Strategic Plan, published in October 2016, establishes a set of objectives for federally funded AI R&D, including both research within the government as well as federally funded research conducted outside the government, such as in academia. One of the goals is to make long-term investments in AI research to drive discovery and to enable the United States to remain a world leader in AI. Another goal is to conduct research to understand and address the ethical, legal, and social implications of AI, as well as to ensure the safety and security of AI systems. In addition, a recommendation is to study the national landscape for creating and sustaining a healthy AI R&D workforce.

---

[4] Intelligent Robotics and Autonomous Systems (IRAS) Interagency Working Group (IWG), https://www.nitrd.gov/nitrdgroups/index.php?title=IRAS

Much basic research on AI and AS is sponsored by DoD, and is conducted at universities. In the 1980s, Office of Naval Research (ONR) re-energized AI research by investing heavily in artificial neural networks (ANN), which evolved into DL. Today, DL, including CNN, are the most powerful machine learning tools in AI, particularly for object recognition in images. ONR also has a very large research program on AS. ONR is interested in exploring autonomous robots for firefighting and maintenance on ships and in exploring autonomous unmanned submarines for ocean mine hunting. DARPA is also investing heavily in AI and AS. DoD AI/AS development is mostly carried out in the commercial sector, teaming with universities to speed up the transition from basic research to technology development. In a recently completed Autonomous Research Pilot Initiative (ARPI) program conducted at DoD Research Laboratories, semi-autonomous robots were demonstrated to be capable of collaborating with humans in accomplishing military missions such as surveillance and reconnaissance.[5]

Although the United States remains the innovative leader, it is not alone in AI/AS R&D. Intense international competition for AI supremacy exists. Japan is well-known for using fuzzy logic AI to control national railways. The robotics industry is more important in Japan than in any other country; it employs more than 250,000 industrial robotic devices. Established in 1982, the European Coordinating Committee for Artificial Intelligence (ECCAI) coordinates AI R&D in Europe and promotes AI study, research, and application. In addition to programming AI to operate in conventional digital computers, the European Union is also investing heavily in building artificial brains (i.e., neuromorphic computing) to implement AI. Chinese technology companies such as Baidu are also investing heavily in AI technology. China wants to be an AI world leader due to AI's strategic importance for national security and economic growth.

China also has a massive market for AI adoption. *MIT Technology Review* cited Baidu's AI work in speech recognition for its low error rate. Prodigious venture capitalist investment in China is sustaining AI R&D in startup companies. A dozen banking and insurance companies in China, as well as state lenders like the China Merchant Bank, are providing funding to help companies develop AI software and to provide efficient services. Facing such competitors, the United States cannot afford to underfund AI R&D or it will be left behind in AI technology innovation.

Expectations are for the domestic AI market to grow rapidly, reaching $70 billion by 2020 according to Merrill Lynch/Bank of America. Access to foreign markets will depend on harmonizing international safety standards and regulations. Similarly, there is intense international competition in AS R&D, due to the significant impact AS will have on the economy. In the United States, in addition to Google and Tesla, major auto manufacturers such as GM and Ford are heavily investing in R&D on self-driving cars. Internationally,

---

[5] https://www.acq.osd.mil/chieftechnologist/cto/cto_arpi.html

BMW, Mercedes, Volvo, Toyota, and Honda are also investing heavily on R&D to speed the development of self-driving cars. Even cities like Hong Kong have introduced self-driving cars and taxis. The first company and country to develop a fully autonomous vehicle will reap the tremendous economic and social benefits.

3. Workforce Development

If you studied computer science or computing engineering, you may already be working on AI or AS. If you are not in an AI/AS-related field, you might wonder, will my job be eliminated? Will a robot replace me? According to one technologist[6], AI/AS will replace 50% of all jobs in the next decade. This claim, however, is not backed up by any rigorous economic or workforce study. A rigorous study of the impact of automation in manufacturing, agriculture, and utilities across 17 countries found that robots, instead of displacing humans, only reduced the hours of lower-skilled workers, but it did not decrease the total hours worked by humans. A Kinsey report concludes that 60 percent of all occupations globally have at least 30 percent of constituent activities that could be automated, while less than 5 percent of all occupations can be completely automated using current technologies, A recent Gartner report concludes that AI will actually create more jobs than it eliminates by 2020.[7] In other words, automation may affect the kind of work humans do, and it may result in structural changes in the economy rather than a net loss of jobs. The most likely jobs that AI will replace are telemarketers, bookkeeping clerks, receptionists, proofreaders, and salespeople. AS will also likely replace couriers, delivery people, and taxi drivers.

No doubt, AI and AS developments will create jobs requiring knowledge and skills in science, technology, engineering, and mathematics (STEM). AI's extraordinary growth creates demand for knowledgeable personnel in AI, as well as related fields. Numerous public and private organizations have recommended increasing student and workforce knowledge in AI expertise. In addition to AI knowledge, development of AS will require studies in multidisciplinary fields such as computer science, systems engineering, and robotic control. DoD faces a particularly difficult challenge in STEM education for U.S. students. Research scientists and engineers working in the military are almost always required to have security clearance and to be U.S. citizens. A large percentage of graduate students in computer science and electrical engineering are foreign nationals, including many from China, who cannot participate in sensitive DoD programs. Hence, efforts must be made to increase the number of U.S. students studying AI and AS.

---

[6] Sophia Yan, "Artificial intelligence will replace half of all jobs in the next decade, says widely followed technologist", CNBC, April 27, 2017. https://www.cnbc.com/2017/04/27/kai-fu-lee-robots-will-replace-half-of-all-jobs.html

[7] Helen Poitevin et al., *Predicts 2018: AI and the Future of Work*, November 28, 2017. Gartner ID: G00342326.

Internationally, fierce competition exists for engineering talent with AI/AS expertise, leading to the threat of the U.S. workforce losing its AI/AS talent. Addressing workforce needs will help to ensure the United States maintains its technological competitiveness internationally and has a workforce that continues to acquire relevant future AI/AS skills. Private sector and academic stakeholders hold a clear consensus that effectively governing AI and AI-related technologies requires a level of technical expertise that the federal government does not currently possess. Effective governance requires additional experts able to understand and analyze the interactions among AI technologies, programmatic objectives, and overall societal values. Creating a public policy and a legal and regulatory environment — allowing innovation to flourish, while protecting the public — requires significant AI/AS technical expertise.

4. Data Security, Privacy, and Transparency

The U.S. Government collects a large quantity of data of all sorts. Increasingly, AI techniques, including machine learning, are used to analyze the data used to make decisions regarding governmental services. As an example, social security cost-of-living increases are based on the previous year's consumer price index. The quality of the data collected for this and other services is extremely important in making the right decisions affecting millions of American citizens. If the machine collects the data without any human verification, how is the machine going to validate the data? In the private sector, AI techniques such as ANN are often used for credit card fraud detections. For the U.S. Government, such AI techniques could be used for fraud detections when someone steals a social security number and fraudulently files for benefits. These and other potential applications of AI to government services highlight the need for government policies ensuring security, privacy, and transparency in such applications and their associated data collections.

Data security obviously is of great concern. Every day brings news about cyber security breaches. Several years ago, the Office of Personnel Management database with all the federal government employee personal information was broken into. The U.S. Government must convince the public that every effort is being made to ensure the data is secured. Privacy is another issue of concern. In spite of the large quantity of personal information the government collects, the public has the right to privacy.

The recent U.S. Government report entitled *Preparing For The Future Of Artificial Intelligence* calls for increasing use of AI in government to improve services and benefit the American people.[8] This report explicitly recommends that *the Federal Government*

[8] Executive Office of the President, National Science and Technology Council, Committee on Technology. *Using AI in Government to Improve Services and Benefit the American People*, October 2016, 15–16.

should explore ways to improve the capacity of key agencies to apply AI to their missions. This includes improving human-machine interfaces through the use of natural language processing and speech recognition when the public uses the telephone to inquire about social services. As an example, current wait-time for calls to the Social Security Administration are exceedingly long (as much as an hour). AI could significantly shorten the wait-time by efficient human-machine interface. Concerns about the use of AI in the delivery of governmental services are not limited to the U.S. Government. The Government of Canada is formulating policies for the responsible development of AI to deliver services more timely and efficiently than ever before, including speedier processing of social benefits.

AI can make decisions based on machine learning and analysis of the data, just like humans make decisions. AI techniques, especially DL techniques, have a tendency to take a "black-box approach" without explanation of why and how the decision is made. In other words, data is simply provided to the machine, and the AI algorithm makes the decision. For example, a citizen may apply for social security benefit, and the AI algorithm will determine approval or disapproval based on the information provided to it. Transparency of the analysis and decision is very important for the public's acceptance of AI technology. Transparency means that the decisions made by the machine must be explainable in terms such that the public can understand how the decision is made. The public has the right to question why and how the decision is made. The DARPA Explainable AI program is beginning to address the problem of explaining machine learning decisions.

5. Safety, Risks, and Cybersecurity

Consumer adoption of AI and AS depends on the public's trust; their opinions about technology safety; and how it affects privacy, employment rates, and the economy. These factors will drive the AI and AS policymaking agenda. The technology can both capture the imagination, as it does with self-driving cars, and inspire fear, as generated in many science-fiction entertainment scenarios. Thus, the public must understand the difference between AI reality and science fiction if it is to accept and trust AI applications as an integral part of modern living. Acceptance and trust will develop only to the extent that the federal government addresses AI and AS safety and privacy concerns and helps individuals displaced by AI and AS to obtain alternative quality employment.

We are currently seeing a whole new generation of unforeseen autonomous system development, such as AI-controlled machines able to learn and make entirely independent decisions. Drones and self-driving cars are prime examples. Naturally, questions exist, such as "Are these autonomous machines safe?" and "If something goes wrong, is there a 'kill switch' for humans to disable the machine?"

Despite the benefits of technology removing the need for human effort, AI and AS also present many challenging safety issues that could negatively impact economic

prosperity and national security (directly or indirectly). AI-enabled AS system failures, such as the fatal crash of a Tesla vehicle operating in a partially automated driving mode, could set the technology back many years. In October 2017, an unmanned drone flying at about 1,500 feet (far higher than the allowed 300 feet) and about 1.9 miles from the airport (much closer than the allowed 3.4 miles) slammed into a passenger plane over Quebec City airport. The collision caused minor damage to the plane and, fortunately, no injuries to the passengers and crew members.

These risks are not limited to robotic AS. AI-enabled expert systems for data analysis also present potential risks for wrong decision-making based on faulty analysis. For example, an AI algorithm may make an incorrect medical diagnosis resulting in prescription treatment and undue patient suffering or even death. An incorrect stock trading decision may cause an investor to lose millions of dollars. The public must be aware of, and willing to accept, these risks in AI-enabled data analysis systems.

Achieving transparency in designing and using a system remains a challenge and a hurdle to adopting AI technology. We increasingly use AI technology to control critical infrastructures ranging from the financial sector to the electric grid. AI safety considerations and risks vary considerably across domains, and federal agencies must establish an understanding of, and develop guidance to promote, responsible adoption of these technologies for public and industry acceptance. In addition, a clear understanding of the safety challenges and risks, and the extent of potential threats and vulnerabilities is critical to formulating public policies and regulations.

AI and AS have powerful implications for national/homeland security. The military is keenly aware of the potential to increase its capabilities while reducing U.S. casualties. In July 2017, DoD launched Project Maven to deploy AI to the war zone by year's end. Project Maven focuses on computer vision – an aspect of machine learning and deep learning that autonomously extracts objects of interest from moving or still imagery. Biologically inspired neural networks are used in this process, and deep learning is defined as the application of such neural networks to learning tasks. Then, in early 2018, DoD released the AI strategy, including a roadmap, to speed up the applications of commercial AI breakthroughs to the military. The AI roadmap includes development of a workforce that understands AI and can implement it to improve DoD's command and control systems, conduct intelligence analysis, and enable effective human-machine collaboration to accomplish a given mission. The Pentagon has created a new Joint Artificial Intelligence Center (JAIC) that will have oversight over almost all service and defense agency AI efforts.

DoD Directive 3000.09, dated November 21, 2012, establishes DoD policy and assigns responsibilities for the development and use of autonomous and semi-autonomous functions in weapon systems, including manned and unmanned platforms. The U.S. has used semi-autonomous weapon systems, such as drones, for at least two decades. When

incorporating autonomous weapon systems into defense planning, it is important to ensure that the operations are always acting in accordance with international humanitarian law. The United States is actively participating at the highest level in ongoing international discussions on lethal autonomous weapon systems. However, using AI to support autonomous weapons is controversial. In 2015, more than 1,000 AI researchers released an open letter calling for a ban on lethal autonomous weapons. That letter was followed in 2017 by an open letter signed by 116 founders of robotics and AI companies from 26 countries that urges the United Nations to address the challenges of lethal autonomous weapons and ban their use internationally. With the rapid developments, AI may cause an international arms race and global instability, especially when combined with robotics. Just as with other powerful weapons, certain international agreements may be necessary to appropriately shape development and use.

In theory, AI and AS technology for the military is not different from that for the civilian sector. For example, self-driving cars can be used in the battlefield to perform reconnaissance, and drones for package delivery can also be used for surveillance or strike in the battlefield. The difference is in the purpose and payload and use of the technology. In practice, however, systems enabled by AI and AS operate differently in civilian and military domains. The rules and laws in military are different from those in the civilian sector. Issues such as safety and trust are considered differently in military and civilian systems, and they should be addressed separately. In the civilian sector, domestic deployed AI-enabled systems used for public safety purposes, may be vulnerable to unauthorized access by people or governments harboring malicious intent, both foreign and domestic.

AI has the potential to affect just about everything in our society. High-profile examples of AI use includes autonomous vehicles (such as drones and self-driving cars), medical diagnosis, creating art (such as poetry), proving mathematical theorems, playing games (such as Chess or Go), search engines (such as Google search), online assistants (such as Siri), image recognition in photographs, spam filtering, predicting judicial decisions, and targeting online advertisements. Just like other computer systems, cybersecurity in AI-enabled autonomous vehicles is a significant issue of concern, which caused the U.S. House of Representatives to enact legislation requiring automakers to report on a cybersecurity plan for self-driving cars.

6. Trust, Ethics, Laws, and Regulations

For acceptance of the rapidly developing AI and AS technology, society must have confidence that these systems can be trusted to make the right decisions. Trust is important in civilian systems and particularly in military systems. Extensive test and evaluation must be conducted to demonstrate that the AI-enabled AS can be trusted. Since AI and AS systems can make decisions entirely on their own, it is very difficult to test and evaluate the systems under all situations. Transparency of the decisions made by the AI algorithm

would help to explain why and how the decisions are made. In the military, decision support systems in the past were developed to assist decision makers on the battlefield to gain situational awareness. With AI technology, the support systems can now be decision systems, making decisions in place of the decision maker.

The rapid pace of development in AI and AS is causing policymakers to question whether the existing laws and regulations are applicable. For example, if a self-driving car hits a pedestrian, who is liable when there is no driver in the car? Similarly if an AI-enabled medical diagnostic system makes the wrong diagnoses, can the computer be sued for malpractice? The legislative agenda in the United States is still developing. The recently enacted SAFE DRIVE Act, H.R. 3388, is just the beginning of the effort to regulate the safety and cybersecurity of autonomous vehicles. The Congressional Robotics Caucus and the AI Caucus are intensely examining the applicability of existing laws and regulations.

The military faces a different set of issues. As far back as 1969, General William Westmoreland, then U.S. Army Chief of Staff, postulated the possibility of an automated battlefield consisting of no civilian or military personnel. The invading force would deploy robotic vehicles and weaponry, while the defending force would counter with similar automated weaponry. Human observers representing both sides would monitor the action from a safe distance, directing and overriding the actions of the robots. It seemed a little far-fetched at the time, but it is increasingly becoming a possibility. Today, UAVs and drones are widely deployed in the battlefield for surveillance and reconnaissance and strike against high-value targets. The Navy plans to deploy a submarine drone squadron by 2020, including the Large Displacement Unmanned Underwater Vehicle (LDUUV), which is highly autonomous with minimal human intervention. The dilemma facing automated warfare was noted in an Air Force treatise, *Unmanned Aircraft Systems Flight Plan 2009–2047*. The dilemma goes as follows: authorizing a machine to make lethal combat decisions is contingent upon political and military leaders resolving legal and ethical questions. These include the appropriateness of machines having this ability, under what circumstances they should be employed, where responsibilities for mistakes lie, and what limitations should be placed upon the autonomy of such systems. In his book, *Army of None: Autonomous Weapons and the Future of War*, Paul Scharre offers a sobering perspective on the automated battlefields. In his book, *Governing Lethal Behavior in Autonomous Robots*, Ronald Arkin argues that it is possible to embed ethics into autonomous military robots and drones to adhere to international humanitarian law and the rules of engagement.

Ethical issues are also important in civilian AI and AS systems. For example, when a self-driving car is faced with a decision to avoid hitting a pedestrian or a parked car, what would the decision be? For humans that decision is not difficult. We value human life more than the car. Will a machine learn the same value, and make the right decision? Recently the National Science Foundation awarded a three-year grant to Professor Nicholas Evans

at University of Massachusetts Lowell to construct ethical answers to questions about autonomous vehicles, translate them into decision-making AI algorithms for the vehicles, and then test the public health effects of those AI algorithms under different risk scenarios using computer modeling. The Berkman Klein Center and MIT Media Lab, with support from the Ethics and Governance of AI Fund, have formed an interdisciplinary team to conduct research to ensure that AI develops in a way that is ethical, accountable, and advances the public interest.

Some researchers and scientists worry that AI will spin out of control. Responsible development of AI and AS will be necessary to convince the public that these fears are unfounded and are based more on Hollywood movies than AI reality. Educating the public about the AI facts and fantasies is critically important to realizing the benefits of AI/AS technology. As leading AI researcher Rodney Brooks writes, "I think it is a mistake to be worrying about us developing malevolent AI anytime in the next few hundred years. I think the worry stems from a fundamental error in not distinguishing the difference between the very real recent advances in a particular aspect of AI, and the enormity and complexity of building sentient volitional intelligence."[9]

7. Summary

AI and AS will significantly affect every aspect of society, economy, commerce, manufacturing, and national security. With AI, computer systems will be able to perform tasks normally requiring human intelligence, such as visual perception, speech recognition, learning, decision-making, and natural language processing. Recent successful AI demonstrations have stirred the public's imagination, such as IBM's Watson competing against a human Jeopardy champion and winning and Google besting a human champion in the ancient Chinese board game Go. Popular demonstrations of self-driving cars have contributed to renewed interest and excitement about autonomous robots. AS integrating AI and multiple sensors will be able to operate entirely on their own without intervention of a human operator. AI and AS will profoundly change the way we live, work, play, and fight wars. However, development of AI and AS is not without risks, including those to safety, privacy, and cybersecurity. Much research and development is yet to be done to realize the full potential and benefits of AI and AS.

Federal investment in AI and AS research and development is essential to stimulate the economy, maintain U.S. competitiveness, create high-value jobs, and improve government services. The recent acceleration of successful AI applications makes it timely to focus R&D investment in more capable AI systems, and also to maximize societal benefits and mitigate any associated risks. Given the potentially significant societal

---

[9] Rodney Brooks, "Artificial Intelligence Is a Tool Not a Threat", *Rethink Robotics* (blog), November 10, 2014, http://www.rethinkrobotics.com/blog/artificial-intelligence-tool-threat.

benefits from AI/AS R&D investment, the federal government should greatly increase IRAS R&D funding. The goals are to make long-term investments in AI research to drive discovery, to enable the U.S. to remain a world leader in AI, and to conduct research to understand and address the ethical, legal, and social implications of AI, as well as to ensure the safety and security of AI systems. Due to intense international competition for AI/AS supremacy, the United States cannot afford to underfund AI R&D or it will be left behind in AI/AS technology innovation.

Concerns about loss of jobs are expected whenever a new technology is deployed. Historically, introducing advanced technologies has always created newer, but different jobs. However, the ratio of job creation to job loss is difficult to predict. Many new jobs required by broadening application of AI technologies will require knowledge and skill in STEM. Other new jobs in manufacture, distribution, sales, maintenance, repair, and recycling of AS and their components may only require routine retraining of existing workers in those fields. AI's extraordinary growth has already created high demand for knowledgeable personnel in AI, as well as related fields.

Research scientists and engineers working in the military almost always require security clearances and U.S. citizenship. Hence, meeting DoD AI workforces needs requires efforts to increase the number of U.S. students studying AI and AS, especially in advanced degree programs. Effective governance requires technical expertise in the federal government to understand and analyze the interactions among AI and AS technologies, programmatic objectives, and overall societal values.

The U.S. Government, collects a large quantity of data of all sorts. Increasingly AI techniques including machine learning are used to analyze the data to make decisions regarding governmental services. Data security and privacy are of great concern. The U.S. Government must convince the public that every effort is made to ensure their data is secured. Transparency of the analysis and decision is very important for the public's acceptance of the AI technology. Consumer adoption of AI and AS depends on the public's trust and their opinions about technology safety and how it affects privacy, employment rates, and the economy. These factors will drive the AI and AS policymaking agenda. The technology can both capture the imagination, as it does with self-driving cars, and inspire fear, such as generated in many science fiction entertainment scenarios. The public must understand the difference between AI reality and science fiction if it is to accept and trust AI applications as an integral part of modern living.

AI and AS have powerful implications for national/homeland security. In theory, AI and AS technology for the military is not different from that of the civilian sector. In practice, however, systems enabled by AI and AS operate differently in civilian and military domains. The rules and laws in military are different from those in the civilian sector. Trust is important in civilian systems and particularly in military systems. In the civilian sector, domestic deployed AI-enabled systems used for public safety purposes may

be vulnerable to unauthorized access by people or governments harboring malicious intent, both foreign and domestic.

Despite the potential benefits of technology removing the need for human effort, AI and AS also present many challenging safety issues that could negatively affect economic prosperity and national security (directly or indirectly). AI-enabled AS system failures, such as the fatal crash of a Tesla vehicle operating in a partially automated driving mode, could set the technology back many years. These risks are not limited to robotic AS. AI-enabled expert systems for data analysis also present potential risks for wrong decision-making due to faulty analyses. The public must be aware of and willing to accept these risks in AI-enabled data analysis systems. We increasingly use AI technology to control critical infrastructures, ranging from the financial sector to the electric grid. AI safety considerations and risks vary considerably across domains, and federal agencies must establish an understanding of and develop guidance to promote responsible adoption of these technologies without stifling innovation. A clear understanding of the safety challenges and risks, the extent of potential threats, and the vulnerabilities is critical to formulate effective public policies and regulations.

Further readings

1. *Issues in Autonomous Vehicle Deployment,* Congressional Research Services Report 7-5700, September 19, 2017.

2. *Shaping Robotics Policy for the 21st Century,* American Association for the Advancement of Science Report, September 28, 2017.

3. *Summer Study on Autonomy,* Defense Science Board, June 2016.

4. *The National Artificial Intelligence Research and Development Strategic Plan,* National Science and Technology Council, October 2016.

5. Artificial Intelligence and Life in 2030, *One Hundred Year Study on Artificial Intelligence* (Stanford: Stanford University, 2016). http://ai100.stanford.edu/2016-report.

6. *DoD Directive 3000.09 Autonomy in Weapon Systems*, USD(P), November 21, 2012,

7. C.F. Barnaby, "Automated Warfare is on the Way: What Are the Consequences?" *Law/International Law/Law of War and Self Defense,* May 23, 2017.

8. Paul Scharre, *Army of None: Autonomous Weapons and the Future of War* (New York: W.W. Norton, 2018).

9. Ronald Arkin, *Governing Lethal Behavior in Autonomous Robots* (Boca Raton: CRC Press, 2009).

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE (DD-MM-YY)	2. REPORT TYPE	3. DATES COVERED (From – To)
00-06-18	Non-Standard	

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Megatrend Issues in Artificial Intelligence and Autonomous Systems	HQ0034-14-D-0001
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBERS

6. AUTHOR(S)	5d. PROJECT NUMBER
Clifford G. Lau, Brian A. Haugh	C5190
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESSES	8. PERFORMING ORGANIZATION REPORT NUMBER
Institute for Defense Analyses 4850 Mark Center Drive Alexandria, VA 22311-1882	NS D-9144

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR'S / MONITOR'S ACRONYM
Insitute for Defense Analyses	IDA
	11. SPONSOR'S / MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

Project Leader: Clifford G. Lau

14. ABSTRACT

Megatrends are sustained developments that fundamentally impact business, economy, society, cultures, and personal lives. Recent advances in Artificial Intelligence (AI) will enable Autonomous Systems (AS), with far-reaching implications in both the civilian sector and defense. AI-enabled robots will perform difficult and dangerous tasks that require human-like intelligence. Self-driving cars will revolutionize automobile transportation systems and reduce traffic fatalities. Big-data analytics using AI techniques will make human-like decisions to improve governmental social services, health care, criminal justice and the environment. AI-enabled autonomous robotic soldiers, aerial drones, and underwater and land vehicles will perform military missions. These revolutionary technological advances will have significant impacts on the economy, military, and society. We are seeing a whole new generation of AI and AS that will change, in unforseen ways, how we live, work, play, and fight wars. However, for the public and military to adopt AI and AS, society and military must have confidence that these systems are trustworthy and safe. A number of important issues are awaiting policymakers, including research and development, workforce development, safety, cybersecurity, ethics, regulations, and automated warfare.

15. SUBJECT TERMS

Artificial intelligence; autonomous systems; cyber security; safety; ethics; regulations; automated warfare

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Insitute for Defense Analyses
Unclassified	Unclassified	Unclassified	Unlimited	18	19b. TELEPHONE NUMBER (Include Area Code)