Artificial Intelligence – Emerging Opportunities, Challenges, and Implications for Policy and Research (GAO Report)

Reports findings from a consortium of Artificial Intelligence experts and stakeholders regarding the technology's evolution and development over time; future opportunities, challenges, and risks of implementing the technology; and new research priorities for regulation.

Updated last **July 24, 2018**
for the 06/26/2018 version of the report.

WHAT IT DOES

In March 2018, the Government Accountability Office (GAO) published "Artificial Intelligence – Emerging Opportunities, Challenges, and Implications" following its July 2017 forum on Artificial Intelligence (AI). This forum convened AI experts and stakeholders from industry, academia, government, and nonprofits to consider the impacts and policy implications of AI in the cybersecurity, transportation, criminal justice, and financial sectors. Following the GAO's report, a summary testimony of the Forum's findings was presented before the Subcommittees on Research and Technology, and Energy within the House Committee on Science, Space, and Technology.

The GAO's report to Congress provides a brief overview of the forums findings by focusing on the following three topics:

1. AI's definition and evolution over time;
2. Future opportunities, challenges, and risks of AI;
3. Research priorities and policy implications of AI advancements.

AI's Definition and Evolution Over Time

Since AI's first inception at Dartmouth College in 1956, the technology's capabilities and attempts to define them have grown exponentially throughout the years. In lieu of establishing a unified definition of AI, the GAO report instead cites a description of the technology's evolution in three waves, each divided by the technology's ability to perceive, learn, abstract, and reason though challenges and data presented to it.

- In the first wave of AI, the technology can automate human pre-programed actions or services using data that humans have made readable to the AI system (e.g. logistical planners and tax preparation services).
- In the second wave of AI, the technology can translate visual, auditory, and semantic information from its surroundings into a format that the AI system can perceive and respond through processes known as machine-learning with limited human oversight.
- In the third wave of AI, which the GAO forum indicates that we are just entering in to, the technology can adapt its operations to new contexts and objectives without human oversight while also explaining how it has adapted meet these changes.

Future Opportunities and Challenges of AI

To shed light on the spectrum of opportunities, challenges, and risks of AI, the GAO forum considered the impacts of AI in the cybersecurity, transportation, criminal justice, and financial sectors. Overall, the GAO report finds that while AI has the capability to improve safety, justice, and security in these sectors, the technology can also undermine these advances if they used maliciously or

without proper oversight. The GAO's report also provides the following summary of AI's opportunities and challenges:

- Opportunities:
  - Improved economic outcomes and productivity: like other technological advancements in the past, AI will improve the rate and efficiency of production. However, the report also mentions that measuring AI's impact will be difficult and there are currently no available mechanisms to accurately measure its impact.
  - Improved or assisted human decision making: AI enables its users to integrate and discover trends or abnormalities hidden within enormous and diversified datasets. Policymakers can use AI systems to create data-driven policy, though validation and potential the programmed-bias of such systems is not yet well understood.
  - Improved problem solving: current progress in AI research promises increasing applications of the technology to society's challenges while also minimizing regulatory oversight burdens to the Government and those being regulated.
- Challenges:
  - Barriers to data collection and sharing: AI systems using dissimilar sources of data may face challenges accessing and integrating data from sources that vary in their data's regulatory accessibility, completeness, and overall quality.
  - Limited access to computing resources and human capital: developers, researchers, and implementers in various governmental organizations or agencies may have difficulties obtaining and funding the computing power and talent-intense needs of AI systems.
  - Legal and regulatory hurdles: the rapid advancement and application of AI systems have in some ways outpaced the regulatory framework to govern how and these systems should be used effectively and safely in its numerous applications. New technological expertise within the government will be needed to make sure that policy for AI is up to date and appropriate for the technology.
  - Developing ethical, explainable, and acceptable AI applications: as AI systems enhance, and increasingly surpass, human capabilities, it will be important that the actions and decisions derived from these systems are able to be held as accountable as the human decision-makers they are assisting and/or replacing.

Cross-cutting Policy and Research Considerations for AI

In response to the emerging opportunities and challenges of AI, the GAO Forum identified a list of policy and research areas that ought to be considered across the government where AI systems are used and researched. The Forum's policy and research considerations include:

- Incentivizing improved data collection, sharing, and labeling – To improve the efficiency and safety of AI's application, federal agencies are advised to implement and organize standardized data collection, sharing, and labeling programs, like those implemented by MITRE, the National Institute of Standards and Technology, and the Office of Science and Technology Policy's Subcommittee on Machine Learning and Artificial Intelligence, that protect the privacy and intellectual property of contributors and allow for more accurate outcomes for the technology's use.
- Improving AI safety and security – New regulatory standards are called for to balance the costs and liabilities of cybersecurity and AI-use are shared more equitably among AI system users, developers, and manufacturers.
- Updating the current regulatory framework for AI – the capabilities and nature of AI systems undermine many current regulatory approaches to privacy, liability, and evaluation where AI systems are being implemented. Federal agencies will need to explore new regulatory approaches while cultivating AI expertise to vigilantly evaluate and improve AI regulation as the technology evolves.
- Defining and assessing acceptable risks and ethics decision making for AI – Federal agencies using AI systems will need to create standardized benchmarks of AI system performance, derived from the perspectives of multiple fields of expertise including economics, philosophy, ethics, the law, to test and evaluate AI systems' degree of risk and ethical implementation.
- Establishing regulatory sandboxes – As new regulatory approaches are considered for AI systems, the government needs to develop regulatory "safe havens" to protect participating stakeholders from risk and liability to allow for more robust participation and evaluation of the new approaches.
- Understanding AI's impact on the Nation's employment and establishing improved job training and readiness programs – The federal government will need to establish more comprehensive data collection to better assess the impact of AI systems on individual and overall employment as well as understand what job sectors will need to be retrained and what new jobs skills will

need to be taught.
- Exploring computations ethics and explainable AI – As AI systems are further developed and used in more contexts, the government and regulatory stakeholders will have to remain vigilant of new ethical considerations for AI's use and the technologies that enable it, like machine learning, big data, and high-powered computer systems.

RELEVANT SCIENCE

There is currently no universally agreed-upon definition of artificial intelligence. The term "intelligence" is understood as a measure of a machine's ability to successfully achieve an intended goal. Like humans, machines exhibit varying levels of intelligence subject to the machine's design and training. However, there are different perspectives on how to define and categorize AI.

Most of the progress seen in AI has been considered "narrow," having addressed specific problem domains like playing games, driving cars, or recognizing faces in images. In recent years, AI applications have surpassed human abilities in some narrow tasks, and rapid progress is expected to continue, opening new opportunities in critical areas such as health, education, energy, and the environment. This contrasts with "general" AI, which would replicate intelligent behavior equal to or surpassing human abilities across the full range of cognitive tasks. Experts involved with the National Science and Technology Council (NSTC) Committee on Technology believe that it will take decades before society advances to artificial "general" intelligence.

According to Stanford University's 100-year study of AI, by 2010, advances in three key areas of technology intersected to increase the promise of AI in the US economy:

- Big data: Large quantities of structured and unstructured data amassed daily from e-commerce, business, science, government, and social media. As datasets increase in size and quantity, so too do concerns about data standardization, securitization, and privatization.
  - Standardization: data provided by multiple parties from multiple sources need to be converted to a common format to allow for consistent collaboration and application by researchers and programs.
  - Securitization: sensitive data must be protected from unauthorized access, manipulation, and application of data throughout the computing system where data is used and stored. A common form of ensuring data security is called "Authentication" where authorized users must verify their identity through multiple methods such as providing a password and a generated passcode sent to the user's phone.
  - Privatization: while a subset of data securitization, data privatization relates to efforts to prevent the disclosure of sensitive information contained in the data such as health, financial, and criminal records. Privatization efforts include anonymizing data as well as providing users transparent indication of who will have access to their data for what purposes.
- Quantum and high-performance computing: Greater storage and parallel processing of big data made possible by emerging computing methods.
  - Quantum computing: whereas traditional computers rely on storing and reading information in binary bits, quantum computers make use of new understandings of quantum mechanics that allow information to be read and stored exponentially faster and simultaneously on non-binary quantum bits or "qubits".
  - High-performance computing: while quantum computing can exponentially increase the abilities of single computers, advancement in high-performance computing enables the simultaneous application of multiple sets of computers, called "clusters", to solve problems. Both quantum and high-performance computing allow for faster and more efficient problem solving, however, these new capabilities could also be applied to nefarious uses that will have to be guarded against.
- Machine learning: the basis for many of the recent advances in AI. Machine learning is a method of data analysis that attempts to find structure (or a pattern) within a data set without human intervention. Machine learning systems search through data to look for patterns and adjust program actions accordingly, a process defined as training the system. To perform this process, an algorithm (called a model) is given a training set (or teaching set) of data, which it uses to answer a question. For example, for a driverless car, a programmer could provide a teaching set of images tagged either "pedestrian" or "not pedestrian." The programmer could then show the computer a series of new photos, which it could then categorize as pedestrians or non-pedestrians. Machine learning would then continue to independently add to the teaching set. Every identified image, right or wrong, expands the teaching set, and the program effectively gets "smarter" and better at completing its task over time.

- Machine learning algorithms are often categorized as supervised or unsupervised. In supervised learning, the system is presented with example inputs along with desired outputs, and the system tries to derive a general rule that maps input to outputs. In unsupervised learning, no desired outputs are given, and the system is left to find patterns independently.
- Deep learning is a subfield in machine learning. Unlike traditional machine learning algorithms that are linear, deep learning utilizes multiple units (or neurons) stacked in a hierarchy of increasing complexity and abstraction inspired by the structure of the human brain. Deep learning systems consist of multiple layers and each layer consists of multiple units. Each unit combines a set of input values to produce an output value, which in turn is passed to the other unit downstream. Deep learning enables the recognition of extremely complex, precise patterns in data.

Experimental research in artificial intelligence includes several key areas that mimic human behaviors, including reasoning, knowledge representation, planning, natural language processing, perception, and generalized intelligence:

- Reasoning includes performing sophisticated mental tasks that people can do (e.g., play chess, solve math problems).
- Knowledge representation is information about real-world objects the AI can use to solve various problems. Knowledge in this context is usable information about a domain, and the representation is the form of the knowledge used by the AI.
- Planning and navigation include processes related to how a robot moves from one place to another. This includes identifying safe and efficient paths, dealing with relevant objects (e.g., doors), and manipulating physical objects.
- Natural language processing includes interpreting and delivering audible speech to and from users.
- Perception research includes improving the capability of computer systems to use sensors to detect and perceive data in a manner that replicates humans' use of senses to acquire and synthesize information from the world around them.

Ultimately, success in the discrete AI research domains could be combined to achieve generalized intelligence, or a fully autonomous "thinking" robot with advanced abilities such as emotional intelligence, creativity, intuition, and morality. Such autonomous agents could open new ethical and legal complications that will need to be adequately assessed and planned for. For instance, autonomous agents or programs may, as a product of their autonomy, operate outside the expectations of their creators. In the event that the agent or program's creators have not implemented comprehensive stop gaps, the agent or program may inadvertently cause unintended harm to allies or adversaries. Whether the creators of the agents or programs are liable for any harms, and whether the harms should be given the same status of acts of war, is yet to be determined.

WHY IT MATTERS

As research and development of AI applications in government and society increase, so too does the need for considerations of the unique opportunities and challenges that accompany the use of AI systems. Central to this report's findings, policy plays a key role in determining what kinds of outcomes will be seen in the evolution of this technology and its effect on society. In this report, a clear need is expressed to continue the government's investment but vigilant monitoring of AI's application.

BACKGROUND

This report was created during several other initiatives throughout the government to assess and address the emergence of Artificial Intelligence in society including:

- The May 10, 2018 White House summit titled, "Artificial Intelligence for American Industry", which reviewed what opportunities, challenges, and prospective policies exist for the US to lead the development of Artificial Intelligence research and applications (SciPol First Look available). During this summit, the OSTP also announced the creation of a select committee on Artificial Intelligence to help coordinate the development and use of Artificial Intelligence throughout the Federal Government.
- The Congressional Artificial Intelligence Caucus, convened by Representative John Delaney [D-MD-6], which similarly convened AI experts and stakeholders to discuss emerging trends, issues, and benefits of AI's rapid innovation and implementation throughout society.

The Office of Science and Technology Policy, under the Obama Administration, also convened a series of forums to assess the emergence of Artificial Intelligence and the government's roll in it. The first was "Preparing for the Future of Artificial Intelligence" (SciPol brief available), which surveys the current state of AI, its existing and potential applications, and the questions that progress in AI raises for society and public policy. The second, National Artificial Intelligence Research and Development Strategic Plan (SciPol brief available), prioritizes key federal R&D investments to maximize the benefits of AI technology. The third, "Artificial Intelligence, Automation and the Economy" (SciPol brief available) recommends policy responses to the projected effects of artificial intelligence-driven automation on the US job market and economy.

ENDORSEMENTS & OPPOSITION

At present, there have not been any publicly reported endorsements of or opposition to this report.

RELATED POLICIES

HR 6090 - Artificial Intelligence Reporting Act of 2018: Requires the Machine Learning and Artificial Intelligence Subgroup of the National Science and Technology Council to report current Federal AI strategies and uses to Congress (SciPol First Look Available).

HR 4625 / S 2217 – Fundamentally Understanding the Usability and Realistic Evolution of Artificial Intelligence Act of 2017: Requires the Department of Commerce to establish the Federal Advisory Committee on the Development and Implementation of Artificial Intelligence (SciPol Brief Available).

PRIMARY AUTHOR

Scott "Esko" Brummel, MA in Bioethics and Science Policy

RECOMMENDED CITATION

Duke SciPol, "Artificial Intelligence – Emerging Opportunities, Challenges, and Implications for Policy and Research (GAO Report)" available at http://scipol.duke.edu/content/gao-report-artificial-intelligence-%E2%80%93-emerging-opportunities-challenges-and-implications (07/24/2018).

LICENSE