



Uploaded to the VFC Website

▶▶▶ 2019 ◀◀◀

This Document has been provided to you courtesy of Veterans-For-Change!

Feel free to pass to any veteran who might be able to use this information!

For thousands more files like this and hundreds of links to useful information, and hundreds of "Frequently Asked Questions, please go to:

[Veterans-For-Change](#)

If Veterans don't help Veterans, who will?

Note:

VFC is not liable for source information in this document, it is merely provided as a courtesy to our members & subscribers.



Proceedings of a Workshop

IN BRIEF

July 2018

Artificial Intelligence and Machine Learning to Accelerate Translational Research

Proceedings of a Workshop—in Brief

The big data revolution, accompanied by the development and deployment of wearable medical devices and mobile health applications, has enabled the biomedical community to apply artificial intelligence (AI) and machine learning algorithms to vast amounts of data. This shift has created new research opportunities in predictive analytics, precision medicine, virtual diagnosis, patient monitoring, and drug discovery and delivery, which has garnered the interests of government, academic, and industry researchers alike and is already putting new tools in the hands of practitioners.

This boom in digital health opportunities has also raised numerous questions concerning the future of biomedical research and healthcare practices. How reliable are deployed AI-driven diagnostic tools, and what is the impact of these tools on doctors and patients? How vulnerable are algorithms to bias and unfairness? How can research improve the process of detecting unfairness in machine learning algorithms? How are other fields simultaneously advancing AI applications? How will academia prepare scientists with the skills to meet the demands of the newly transformed industry? Informed answers to these and other questions require interdisciplinary discussion and collaboration. On **February 13 and 14, 2018**, the Government-University-Industry Research Roundtable convened a workshop to explore these and other questions related to the emerging use of AI and machine learning technologies in translational research.

The keynote speech on February 13 was given by **Oren Etzioni** from the Allen Institute for Artificial Intelligence, who focused his remarks on demystifying AI. “Elon Musk said that with AI, we are summoning the demon,” noted Etzioni in opening. “But we need to separate the science of AI from the science fiction.”

Etzioni explained that in theory, AI is defined as machines doing what humans do. In practice, AI refers to narrower tasks—such as pattern, speech, or image recognition—that were once thought to require human intelligence. Recent progress in computer science shows that machines can complete these narrower tasks as well as—or maybe even better than—humans.

Public demonstrations of AI besting human players at chess and Go (an abstract strategy board game for two players) have blurred the difference between theoretical and practical AI capabilities. “When a computer program, AlphaGo, beat the human Go champion, many people thought that it was showing us the path to full, general artificial intelligence. But board games are black and white; they have discrete moves and defined boards and a very clear win versus lose evaluation. And the amount of data a computer program can draw upon is approximately infinite.”

Etzioni explained that practical AI shines in those kinds of black-and-white situations, but programming AI and machine learning is still 99 percent human work. “The equation for AI success is to take a set of categories (for example, cats and dogs) and an enormous amount of data (that is labeled as to whether it is a cat or a dog), and then feed those two inputs through an algorithm. That produces the models that do the work for us. All three of those elements—categories, data, algorithm—are created through manual labor.”

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The result of these projects is the creation of what AI and computer scientists call “AI savants”—narrow systems that are incredibly good at one thing, but inept at switching from one activity to another, like from Go to chess. “Intelligence is not autonomy,” said Etzioni. “Because people are both intelligent and autonomous, it is easy for us to conflate these two characteristics. But even a self-driving car is not choosing where to go.”

Despite the evidence to suggest that humans are still in control of practical AI systems, policy questions about managing the uncertainty of future AI capabilities and about the technology’s impact on the workforce and implicit bias are persistent and important ones. When asked if there should be a moratorium—on AI research, Etzioni said the country cannot afford one, for two particular reasons.

The first reason is that AI is global. In Etzioni’s opinion, if the United States slows its exploration of AI, it will be overtaken by other countries that are already competing and in some areas beating the United States at this game. China and Russia have both declared the promise of global leadership in the AI space. His second reason is the potential benefits to society that can come from AI.

“Paul Allen founded the Allen Institute for Artificial Intelligence with the goal of exploring AI that can be used for the common good. One thing we work on is using AI to improve scientific search engines. The number of scientific publications is growing exponentially, and we need some way to help researchers and doctors cut through the clutter.” Etzioni believes in this sense that AI is less “artificial intelligence” and more “augmented intelligence”—working shoulder to shoulder with scientists to help them do their job, and perhaps someday this will progress from machine reading to suggesting hypotheses and experiments.

“In a recent piece in *The New York Times* I tried to suggest a pragmatic framework for thinking about AI,” said Etzioni. “For example, if we build AI systems, we should have an ‘off’ switch. When regulating AI and the data and algorithms it employs, we need to think about regulating its applications—cars, toys, robots, medical devices, et cetera—instead of the scientific field itself. We should recognize that people are ultimately responsible when AI systems are in use. ‘My AI did it’ is not an appropriate excuse.” Etzioni also argued two more governing principles for AI: it should not use devices to elicit information from users without their knowledge, and it should not amplify the bias in data. “These systems use data from the past to generate models to predict the future, so if society’s past was racist and sexist, the models will carry that bias into the future and also, for technical reasons, exacerbate it. We have ways to prevent that, which should be implemented widely.”

“One of the trickiest issues is what to do about jobs,” Etzioni said. “Some of my colleagues say we need technical training, and certainly we need more understanding of and exposure to AI. But it is not realistic to think we are going to take coal miners and turn them into data miners.” One suggestion Etzioni put forth was to suggest caregiving as an alternative career area for displaced workers. “Though it is by no means a comprehensive solution, caregiving is a uniquely human job,” Etzioni concluded.

AI APPLICATIONS FOR DRUG DISCOVERY, DIAGNOSIS, AND TREATMENT

The first panel on February 14 explored biomedical applications of AI and its implications for drug discovery, diagnosis, and treatment. The first presentation was given by **Alvin Rajkomar** from Google Brain, a deep learning AI research team at Google. He focused on two areas where AI and machine learning can make a difference in health care, though he noted that there are many more.

The first area he mentioned was computer diagnosis. According to Rajkomar, “Teaching computers to interpret medical images, involving human experts interacting with a computer model, is not a trivial matter, but medical communities are already adopting the practice. In diabetic retinopathy—one of the most rapidly growing causes of blindness—there are not enough trained ophthalmologists to diagnose all of the cases, especially in developing countries. It is now possible for a machine to screen images of people’s eyes, identify those with signs of disease, and refer those people for examination and treatment by an ophthalmologist. This work was published in the *Journal of the American Medical Association*, and the approach works well—probably better than identification by a human,” said Rajkomar.

Rajkomar mentioned that another area of machine learning applications is care-management decision support—helping doctors make better decisions for patients, catching errors, and performing other assistive functions. “Many patients who enter a hospital will have an adverse reaction to some aspect of hospitalization and treatment. If a machine can predict who is at very high risk, then a team could work proactively with that patient to make sure any adverse reactions do not happen,” said Rajkomar. He described some of Google’s recent work to help predict which patients would have trouble in the hospital. The team developed a software-engineered pipeline that converts data from different health systems into a single format; once all the data from different health systems diagnoses, the

medications a person is taking, lab test results, and what the doctors and nurses wrote about the patient were included in the system, the team time-ordered it from start to finish. They then used all of the data to read these records from start to finish and to make predictions about patients admitted to the hospital—for example, who is likely to die in the hospital and who is likely to be readmitted. The models were able to make predictions more accurately and earlier than standard models.

Rajkomar echoed Etzioni's comments about the importance of identifying data and algorithmic bias. "We take fairness and bias seriously," said Rajkomar. "Machines learn from the data you give them, and sometimes machines learn correlations—for example, involving race—that you do not want them to learn. It is important to ask how these machines could be unfair to people and communities and what the unintended consequences might be."

Jed Pitera from IBM Almaden Research Center spoke next about using AI for drug discovery efforts. Pitera described how Almaden worked with researchers at the University of California, San Francisco to enable ultra-high-throughput microscopy of microbes to understand how the microbes respond to perturbations in their environment or to drugs or pharmaceuticals. "Research is generating an enormous amount of images of cells and cell structures, but the bottleneck is image segmentation—finding out where the cells are in the image, and where structures like organelles are inside the cells. This is currently a very labor-intensive task, and people have been deploying deep-learning algorithms to do this kind of segmentation process. The researchers showed that if this is done in the context of active learning—an iterative process in which the deep learning algorithm proposes a segmentation and then a human expert corrects that segmentation, updating the algorithm—it is only necessary to use about one-sixth of the images that you would need otherwise to train the algorithm to do the segmentation effectively."

In the pharmaceutical space, especially in drug discovery, researchers have used machine learning and deep-learning approaches for mappings between a chemical structure and some physical property, such as water solubility. An advantage to using deep-learning and machine learning in pharmaceuticals is that researchers are often dealing with large, homogenous data sets, so it is easy to generate large amounts of uniformly labeled data useful for training a machine. Material scientists have adopted similar approaches to machine learning models.

Almaden also uses machine learning models as a way to encode relatively subtle aspects of human expertise. Pitera explained how he and his colleagues showed materials scientists about 200 examples of materials and asked them to put each in one of three simple categories: "this material looks good for my application," "this material does not good for my application," or "I don't know." This data was then used to create a simple model that can prioritize conditions in large sets of materials, which can then be used to identify materials that are likely to be interesting or useful. "We have found that this works especially well because people are not good at expressing in a quantitative way some of the deep technical knowledge they have about a particular domain," said Pitera. "The challenge lies in picking up the domain knowledge that lives in the human expert without also picking up the biases."

The next presentation was given by **Sandy Farmer** of Pfizer, whose current role is to look across the research enterprise to see how the company can make better use of its data. "I do not see AI replacing people in the drug discovery area, but rather as a tool to augment—to help scientists do things faster and better and to give them insights," Farmer said.

"The availability of data from clinical biology, cellular biology, and genomics has changed the application of artificial intelligence to biology," he continued. "We are trying to use genomics—nature's experiments on people—to inform on biology, which we can then evaluate progressively backwards using clinical and cellular and in vitro assays, and then do the asset discovery. This can be thought of as a process of reverse translation. That data has always existed, but only now are we pulling it together with all of the integration required."

Pfizer conducted a strategy last year to identify areas where applying AI machine learning could move the industry forward. One area is identifying novel targets and biomarkers using the plethora of data available, and another is helping chemists make different compounds faster. According to Farmer, one problem in the industry has been that chemists generally make the compounds they can make, and not the compounds they should make, because the latter are difficult to identify. Farmer's team asked how they could use the literature and machine learning to give chemists a head start on the compounds they should make.

"The future of the pharmaceutical industry is in making use of human exploratory data," said Farmer. "The only way to streamline access to human exploratory data is if patients allow it, so the solution is patient empowerment. Currently there is no mechanism for this, though there are institutions in the United States and Europe that are exploring models where patients may load their own data into 'sandbox' environments and consent for their data to be used for research."

Farmer stated that the goal for human exploratory data would be to aggregate data across the entire medical ecosystem, and give it to third parties to analyze. The pharmaceutical industry could then use AI to build models or to surface patterns—connecting with the literature—to provide insights into potential benefits to patients. “Industry is only a small part of this effort,” said Farmer. “It is going to take universities, government, and industry—society at large—to make this work.”

Robert F. Murphy of Carnegie Mellon University spoke next, focusing on active machine learning in biomedical research. Murphy opened by explaining that drug screening is typically done by taking an assay that measures a target of interest, and using that assay for hundreds of thousands, or millions, of compounds that could potentially be useful as drugs. Some of those compounds result in the activity the researcher is looking for, but this does not allow the researcher to learn anything about the side effects of those drugs on other targets in the body. According to Murphy, ideally it would be possible to look at all effects of all drugs on all targets, which could facilitate the design of very effective drugs—those that had just the activity wanted with no other side effects. “But that’s a very big matrix of possible experiments, and so we can’t do that exhaustively,” said Murphy.

“We can think of filling out that matrix of possible drugs and possible targets like a game of Battleship,” said Murphy. “We have a grid, and we start with no information, and we want to find the ships (or potential hits on drugs). There are some ships on the grid, but most of the grid is empty. Currently, researchers start in the corner and say A1, A2, A3, A4, etc., but we are not going to win that way.”

He offered an example of how to proceed in a different way, using existing data. Murphy’s research team introduced data on many drugs with many targets from PubChem to an active machine learning system, and did not let it see anything about hits between drugs and targets—just like Battleship. The team tasked the machine to ask for information on the equivalent of A1, A7, B12, etc., and the team reported the data back to it upon its requests. The machine found 57 percent of the hits by doing only 2.5 percent of the PubChem experiments with this active learning method (see Figure 1). In another experiment considering the effects of 96 drugs on 96 different cellular targets, the machine selected which experiments to do based on what it did not know—the researchers were able to get a model that was 92 percent accurate at predicting hits while doing only 28 percent of the possible experiments.

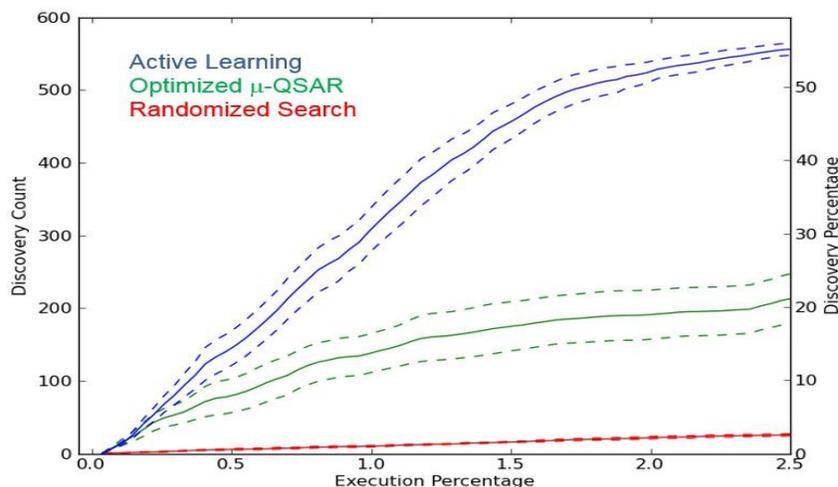


Figure 1 Comparison of effectiveness of three virtual screening models identifying discoveries of “hit” compound effect scores by a programmed active learning machine based on coverage of the experimental space. The active learning model was able to identify 57 percent of the active compounds by covering only 2.5 percent of the matrix.

NOTE: Data source: Kangas, Naik, and Murphy, “Efficient discovery of responses of proteins to compounds using active learning,” *BMC Bioinformatics* 15:143, 2014. SOURCE: Robert Murphy, Carnegie Mellon University.

“There is plenty of room for AI and deep learning to do a lot with existing data in terms of discovering trends and making predictions from data that already exists,” said Murphy. “But going forward, we need to stop trying to simplify things that are inherently complex, and instead embrace that complexity and build models that can handle that complexity.”

EMERGING CROSS-SECTOR AI CHALLENGES AND OPPORTUNITIES

The subsequent panel considered how AI trends and technologies used in medicine were being applied and advanced in other sectors across the research landscape. **Nathan Hodas** from Pacific Northwest National Laboratory (PNNL) gave the first presentation. “All of the areas of PNNL’s work produce a lot of data—in the past it was siloed and mostly ignored because of the lack of analytics capabilities. The lab has found that the answer to big data is machine learning; it is the only technique that can absorb that much data and continue to improve” (see Figure 2).

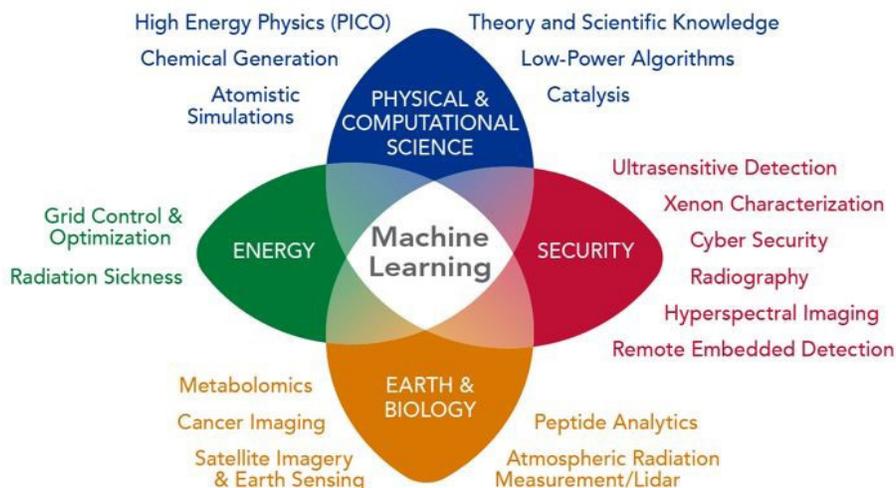


Figure 2 Disciplinary areas of deep learning for scientific discovery at the Pacific Northwest National Laboratory. SOURCE: Nathan Hodas, Pacific Northwest National Laboratory.

Hodas explained machine learning as the process of getting a machine to do what you want through example rather than by programming—a way of teaching a machine to recognize and understand patterns that emerge from data. Research on deep learning and neural networks—extremely effective pattern recognition techniques based on efforts to make computers function like the human brain—has been happening since the 1970s, but according to Hodas this research fell out of favor in the 1990s and early 2000s because of its computational intensity.

Eventually larger amounts of data, better computing, and better software coupled with the preservation of research from those early efforts has led to a resurgence in neural networks. Hodas continued, “Work in just the last two years allows almost anyone who can write a computer program to have a neural network, using three lines of code. A neural network that is trained on images of dogs and cats learns about aspects of images—such as what edges are, what colors are, what an ear looks like—that can be transferred to other problems.”

Hodas explained an example of how deep learning is being applied at PNNL to predict radiation sickness in patients. “The application can predict, based on measurements of the skin taken at the moment of radiation exposure, whether someone has received a harmful or fatal dose of radiation. We have dozens of other areas where if you have the data and the metric it is a very promising avenue for extracting knowledge from the data, but to be successful you also need computing power.”

Hodas does not believe deep learning holds the solution to every problem—though it is a very sharp tool, he conceded that it can do damage if used improperly. “At the national labs we have found an alternative, structured way to bring these tools into disciplines that do know their data. You need a scientific goal, a lot of data, and a domain expert who really understands the data. Lastly, you need a metric—a way to tell that one approach is better than another approach.”

The next presentation was given by **David Gunning** from the Defense Advanced Research Projects Agency (DARPA) on the explainability of AI. He opened by explaining that DARPA characterizes the history of AI research in three waves: the first started in the 1950s and concerned symbolic reasoning, like generating proofs of all of the theorems in *Principia Mathematica*; the second wave concerns statistical learning, which involves developing models

to learn from real data (see Figure 3). Gunning characterized current AI research as part of this second wave. “We are trying to conceptualize what the third wave will be at DARPA,” said Gunning. “We are calling it the explainable or contextual adaptation wave, in which we are trying to combine perception and sophisticated reasoning and abstraction, in order to get close to human-level intelligence.”

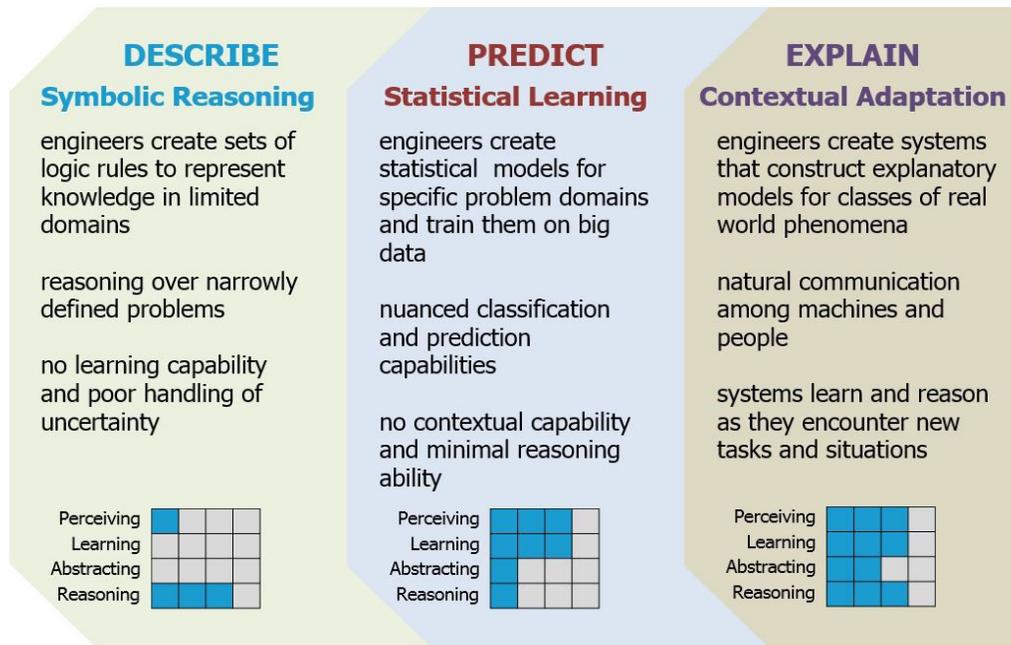


Figure 3 An explanation of the three waves of artificial intelligence research. SOURCE: David Gunning, Defense Advanced Research Projects Agency.

One noteworthy DARPA project is called Big Mechanism, which is developing an AI system that can read research on cell biology, extract enough information out of it to construct a candidate model of cancer pathways in the cell, interact with a human biologist to try to refine the model, and develop a hypothesis and a causal model of what is happening. Another DARPA program, called Lifelong Learning Machines (L2M), enables systems to improve their performance over their lifetimes with machine learning. Current machine learning systems require huge data sets upfront that are used to train the model, but if the data changes, the model does not necessarily keep up. The L2M program is trying to get beyond this limitation so that machine learning programs can adapt models in the real world, and instantaneously.

Gunning described one final program, on Explainable AI. “Right now it is hard for people to understand why AI programs are making the decisions they do,” said Gunning. “If you ask a current AI system how it knows that an image of a cat is a cat, it says “the output is .93 on the cat node,” but it cannot really explain why, which makes it hard for a user to know when he can and cannot trust the technology.” The DARPA program is trying to build a system that allows researchers to get more interpretive features and labels out of the model—an explanation that an application user can understand and use to further improve the model.

The next presentation of the panel was given by **Erik Vinkhuyzen** of Nissan Motor Company, whose job is to look at AI systems from the perspective of the users of systems. Vinkhuyzen described one of many technical problems with modern autonomous vehicles, which concerns their prediction algorithms. These algorithms try to predict where every car, biker, and other element in view will be a few moments from now and use those predictions to plan a safe path. However these prediction algorithms do not take into account people’s intentions. “It is very hard to move from simply perceiving a pedestrian to predicting what that pedestrian is actually going to do. Statistical knowledge about what people are likely to do does not help in a particular driving situation; you had better get it right, there and then.”

Another challenge with predictions is that predictions depend on self; if you are at an intersection and wondering whether someone is going to go first, staying stopped quickly resolves that issue; the other car will go first.

Vinkhuyzen calls these interactions “a very profound puzzle.” Why and how is it that these kinds of interactions run off so effortlessly between people, millions of times a day?

According to Vinkhuyzen, the secret ingredient to making these social interactions run smoothly is trust between human beings—we share, and trust that we share, a common way of perceiving the world. The pedestrian assumes the driver is in control of the car, that the driver can see the road and the pedestrian on it, that the driver knows that pedestrians have the right-of-way in crosswalks, that the driver is not evil or drunk, and that the driver is paying attention. The pedestrian also assumes that the driver knows things about the pedestrian—for example, that the driver can see that the pedestrian intends to cross the road.

“All of these things together have to be in place in order for a pedestrian to cross the road and to do so with confidence. This kind of trust is at the basis of driving safely. Part of the challenge with autonomous vehicles is to drive in such a way that we do not violate these kinds of rules and practices on the road that keep us safe,” said Vinkhuyzen.

Jeff Alstott of the Intelligence Advanced Research Projects Agency (IARPA) gave the final presentation of the panel, and offered remarks about the security of machine learning systems, which he considered an under-recognized problem. Alstott’s main worry is that machine learning systems are not secure and can be fooled. He theorized, “Suppose Alice has trained a neural network and sold it to Bob, who is going to use it in his self-driving car. Bob tests it and finds that the car stops at stop signs, and speeds up at speed limit signs, and so on. But unbeknownst to Bob, during the training of the machine learning system, Alice put in some data in which stop signs with post-it notes on them have been labeled as a speed limit sign. This is a problem, because now the car will not stop at all stop signs, and may speed up at some of them.”

In Alstott’s opinion, transitioning machine learning into settings with safety relevance—such as with autonomous vehicles, critical infrastructure, or national or global security contexts—creates the prospect for new vulnerabilities, in addition to traditional cybersecurity concerns.

“An adversary trained in the system does not necessarily have to exist to have these problems,” he continued. Alstott explained that though the neural networks in use today may have been developed with good intentions, they can still make errors. “Even humans make mistakes with optical illusions,” Alstott said. “The equivalent of optical illusions for machine learning systems is currently rampant. If an adversary is trying to fool your perfectly benign system, they can manipulate the environment, and by doing so they can manipulate the behavior of your system.” Alstott closed by saying this is an area of study that IARPA will continue to direct attention toward.

AI AND INTELLECTUAL PROPERTY

Robin Feldman from the University of California Hastings’ Institute of Innovation Law delivered a presentation on AI and intellectual property. “Obtaining intellectual property protection for AI is uncertain terrain at the moment,” she said. “It is not optimal for either incentivizing innovation or for fostering the type of trust and reliability one might want in this burgeoning field.”

Feldman explained that the critical part of AI rights relates to the software, since a core value flows from the software and what that software produces. “When we speak about intellectual property and AI, we are really speaking about two levels: the first is the intellectual property in the AI program itself; the second is the intellectual property produced by the AI—for example, life-science applications that collect data on a patient in real time and adjust treatment accordingly.”

According to Feldman, copyright law poses a simpler set of legal questions than patent law in the area of AI. Software is clearly protected by copyright law, so the initial AI programs are copyrightable. However, the U.S. copyright office casts some doubt on the copyrightability of programs created by AI itself. The office’s Compendium of Office Practices holds that it will not register works produced by a machine or mere mechanical process that operates randomly or automatically without creative input or intervention by a human author. That language is potentially problematic for AI inventions, though it has not been legally tested in this country.

“There may also be a way to see more room than meets the eye in the position taken by the copyright office,” said Feldman. She described a draft Restatement of Law (nonbinding interpretations of law that are often cited by judges) that says that although a computer program might someday produce an output so divorced from the original program that it raises the question of whether the human creator or user could be regarded as the author, that issue has not arisen yet and therefore current outputs should be copyrightable.

“The more powerful rights, particularly for life sciences inventions, lie in patent law, and there the terrain is decidedly less hospitable,” Feldman explained. “In a series of four cases between 2006 and 2013, the Supreme Court

drastically limited those things that may fall within patentable subject matter. The full quartet of cases makes it increasingly difficult to defend software patents as well as difficult to defend patents on things that are discovered with machine learning.”

Feldman raised several questions to the audience about the future of AI and patent law. Where might we take the law? What pathways might we pursue? How should the United States protect its competitiveness? What types of protection will foster the innovation that secures our place at the head of this revolution? “In finding the right approach, one of the core challenges for protecting AI and its progeny inventions relates to openness,” said Feldman. “Societal acceptance of AI systems may well hinge on a degree of accessibility different from the norm in IP systems. On the simplest level, if people are going to be coaxed into using new devices, we have to establish basic levels of trust.”

Unfortunately, Feldman argued, our patent and copyright systems are not a good match for developing trust, since an inventor does not have to reveal much about the software code to secure a patent or copyright. “Instead of hewing to the current patent or copyright protections as a model for AI-generated work, we might do better to take a page from the protection in drug development,” said Feldman. “In that context, a brand drug company receives 4 to 5 years of exclusivity in exchange for making safety and efficacy data available to the government and competitors—a model that may be more useful as we head into the frontiers of AI.”

AI AND THE FUTURE OF THE WORKFORCE

The day’s final presentation was delivered by **Melvin Greer** from Intel, who spoke on AI in the workforce. Greer opened by stating that by some estimates, thirty percent of the hours worked globally are “automatable.” But the social, ethical, and legal implications of conversations over whether and how jobs can be completed by trained machines are more complex and far-reaching.

“By around 2025, domain knowledge—as opposed to technical knowledge—will be at a premium,” said Greer. “Domain knowledge will encompass the skill set that will drive the ability to stay employed, and people with domain knowledge will be at an advantage in developing AI.” He cited a survey by Gartner that found by 2020, 40 percent of all data science activity will be performed by citizen data scientists who do not have formal training but are responsible for understanding and executing core analytic skills.

“Can we take a truck driver and turn him into a neuromorphic engineer? The answer is yes,” said Greer. “Upskilling workers will be possible in part because those who are developing data science platforms are trying to simplify their platforms to make them easier to use. AI will not just destroy jobs—like the world experienced during the rise of the personal computer, AI will create many jobs and potentially entirely new industries.”

Greer addressed two important questions: How are we going to prepare for future jobs, and what are the best ways to think about managing the changes that AI will bring about? The first question depends on thinking about what will be different about the jobs of the future. According to Greer, “The jobs of the future will emphasize the things that make us unique as humans: skills such as creativity, critical thinking, and emotional intelligence. One job we expect to see in the next 5 years that doesn’t exist today is a personal privacy manager—a person who can adjudicate your digital self.” Greer suggested that human-machine collaboration will also be important in future jobs, especially in industries newly integrating the use of AI applications; AI will perform the repeatable frequent volume tasks, while human workers focus on the new tasks.

To answer the second question, Greer referenced advice from the first Chief Data Scientist of the U.S. Office of Science and Technology Policy, D.J. Patil, who recommended that data scientists need to have technical skills, to be clever, and to be very good storytellers, but they also need to have a grounded understanding of the ethical, legal, and societal implications of their actions. Greer advocated that the challenges evolving in the AI ethics dialogue necessitate emotional intelligence and ethics training in data scientists especially.

“Industry is not waiting to fully investigate and implement AI,” said Greer. “Employers started to understand the workforce implications of changing AI capabilities before workers and policymakers, because the strategic intent of all of this AI is focused on innovation and growth.” Greer also warned against the inequity of perpetuating current data usage practices at the national level without thinking about its impact on citizens, and noted that some are considering a universal basic income to alleviate the economic pressures of job automation while reimbursing citizens for sharing their data. He concluded, “If you do not have data, it is harder to participate in the new AI economy, and this is driving inequity that will require systemic changes to resolve.”

DISCLAIMER: This Proceedings of a Workshop—in Brief was prepared by **Sara Frueh** as a factual summary of what occurred at the meeting. The statements made are those of the rapporteur or individual meeting participants and do not necessarily represent the views of all meeting participants; the planning committee; or the National Academies of Sciences, Engineering, and Medicine.

REVIEWERS: To ensure that it meets institutional standards for quality and objectivity, this Proceedings of a Workshop—in Brief was reviewed by **Stephanie Courchesne-Schlink**, National Institutes of Health, and **Prakash Sesha**, Lockheed Martin Corporation. **Marilyn Baker**, National Academies of Sciences, Engineering, and Medicine, served as the review coordinator.

PLANNING COMMITTEE: **Jeffrey Welser**, IBM Research-Almaden; **Taylor Gilliland**, National Center for Advancing Translational Science at the National Institutes of Health; **Avery Sen**, Toffler Associates.

STAFF: **Susan Sauer Sloan**, Director, GUIRR; **Megan Nicholson**, Program Officer; **Claudette Baylor-Fleming**, Administrative Coordinator; **Cynthia Getner**, Financial Associate.

SPONSORS: This workshop was supported by the Government-University-Industry Research Roundtable Membership, National Institutes of Health, Office of Naval Research, Office of the Director of National Intelligence, and the United States Department of Agriculture.

For additional information regarding the workshop, visit <http://www.nas.edu/guirr>.

Suggested citation: National Academies of Sciences, Engineering, and Medicine. 2018. *Artificial Intelligence and Machine Learning to Accelerate Translational Research: Proceedings of a Workshop—in Brief*. Washington, DC: The National Academies Press. doi: <https://doi.org/10.17226/25197>.

Policy and Global Affairs

The National Academies of
SCIENCES • ENGINEERING • MEDICINE

The nation turns to the National Academies of Sciences, Engineering, and Medicine for independent, objective advice on issues that affect people's lives worldwide.

www.national-academies.org

Copyright 2018 by the National Academy of Sciences. All rights reserved.

Copyright National Academy of Sciences. All rights reserved.