MASSACHUSETTS INSTITUTE OF TECHNOLOGY

STARR FORUM REPORT

May 25, 2018
18-01

Artificial intelligence and national security law: A dangerous nonchalance

Honorable James E Baker, MIT
Former Chief Judge, United States Court of Appeals for the Armed Forces

My topic is artificial intelligence and national security law, with emphasis on law. I know better than to come into MIT and talk about algorithmic formulas. I am not a technologist. I am a bridger. I do not mean I am a civil engineer, but someone who bridges communities. My job is to translate for you, in plain English, why you should care about AI in the national security sphere. And, my job is to explain to policymakers why they should care about national security law as it applies to AI, and care now.

I will start with a few data points, with which I hope to grab your attention. I will then define AI, summarize its evolution, before describing its national security applications and implications. I will close by suggesting a legal framework for moving forward along with two final questions.

Data points

Let's begin with the 2017 Belfer-IARPA study on AI and national security. (IARPA is the intelligence community's version of DARPA.) The study concluded that AI is likely to be as transformative a military technology as aviation and nuclear weapons were before. That should get your attention—as transformative a military technology as aviation and nuclear weapons.

In July 2017, China's State Council, in theory their highest governing body, released an AI plan and strategy calling for China to become the world's leader in AI by 2030. The aim is to reach parity by 2020 with the leading countries, i.e., the United States, to become one of the world's leaders in 2025, and to be the primary AI power in 2030. China has committed $150 billion to this goal. However, as you will see in a moment, it is hard to define AI and it is hard to account for $150 billion across multiple technological fields and subfields.

As many of you know, in 2016, Google's AlphaGo computer, a DeepMind computer at the time, beat the world's best Go player, Lee Sedol. This was viewed, and still is viewed, by many as a Sputnik moment in China. This was in part because Go is a Chinese game and in part because it was a US computer that won. 200 million people watched AlphaGo beat Sedol on television. But not in the US.

The Starr Forum is a public event series hosted by the MIT Center for International Studies through a generous gift from the Starr Foundation of New York.

The talks feature leading thinkers on pressing issues in the world of international relations and US foreign policy.

The **Starr Forum Report** is a new publication that features the talks in print.

Authors in this series are available to the press and policy community.

Contact: Michelle English
english7@mit.edu

Lest Russia feel left out, here are two reassuring quotes from Vladimir Putin. "Whoever becomes the leader in this sphere [AI] will become the ruler of the world." And my personal favorite, "When will it eat us?" I'll explain a little bit more about being eaten in a moment.

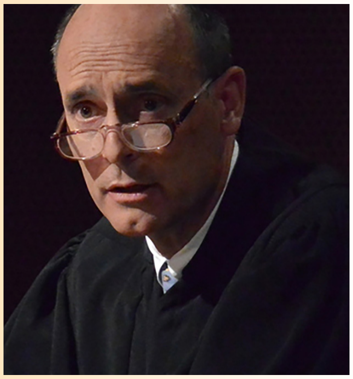
And, of course, you've seen the many warnings from Stephen Hawking and Elon Musk, captured in succinct manner by Hawking, "In short, the rise of powerful AI will be either the best or the worst thing ever to happen to humanity."

AI defined

There are many AI definitions. That is because AI is hard to define. It encompasses many fields and subfields. We also anthropomorphize AI viewing "intelligence" as a human quality and one with emotions. But AI is really programmed machine optimization.

That is why I like the Stanford 100-Year Study definition, which captures much of this nuance. "Artificial intelligence is a science, and a set of computational technologies, that are inspired, but typically operate quite differently from, the way people use their nervous systems and bodies to sense, learn, reason, and take action."

The premise behind AI is that if you can express an idea, a thought, or an action in numeric fashion, you can code that purpose into software and cause a machine to achieve that purpose. One key question for technologists therefore is what exactly are the limits to what you can express numerically and code?



The Honorable James E Baker, a Robert E Wilhelm Fellow at the MIT Center for International Studies, retired from the United States Court of Appeals for the Armed Forces. He joins Syracuse University this fall as a professor at the College of Law, the Maxwell School of Citizenship and Public Affairs, and as director of the Institute for National Security and Counterterroism.

The idea of AI has been around since Bletchley Park, Alan Turing and the Imitation Game. The first AI conference was held at Dartmouth in 1956. The grant proposal indicated that 10 scientists would meet and solve AI. We're still working on that piece, but that sense of confidence—hubris—has returned.

From spring to summer

If you wade into the AI literature, there's a lot of talk about seasons, sort of like New England, there's a recurring discussion about AI winters. That means a pause in the funding for AI, as well as a pause in the progress of AI. But what I'm trying to tell you with quick reference to my data points, and what I would be telling a national security official in the United States government, if I could, is not only is it spring, AI summer is right around the corner. It is going to be a hot summer.

There are several factors for the advent of spring and the prospect of summer.

***The growth of computational capacity found in among other places, super computers and chips*.** For example, an iPhone 5 has 2.7 times the computational capacity of a Cray-2 supercomputer from 1985.

***Data and Big Data*.** In many cases, AI relies on and learns from data. The greater the volume of reliable data the more reliable the AI output. Now think about the Internet of Things, Facebook, Twitter, etc...

The revolution in algorithms and software is the next area that has helped blossom this field.

***The explosion in knowledge and mapping of the brain*.** Many AI people are trying to model neural networks from the brain, some literally by trying to create 3D versions of brains, and others, metaphorically, by creating electronic neural networks.

***The dot.com and Facebook explosion*.** This has led to people pumping money into technology. To turn Napoleon's phrase about a marshal's baton in every soldier's pack, in every dorm room, we now believe, is a billionaire's technological portfolio, if you just come up with the right idea or application.

All these factors come together in what is known as machine learning, which in general, is the use of algorithms and data to train computational machines to detect patterns, classify information, and predict outcomes, often better than humans can in certain spheres.

Having identified these factors, one might pause for a moment to think about China. $150 billion is a lot of money. They have advantages in focus and centralized authority. They have data, lots of it. (Notice that in 2015, they passed a law prohibiting any data involving Chinese persons from being transferred overseas.) I would love to say we have the advantage of market incentive, and they have the disadvantage of communist centralism, but they also have the advantage of market incentive, in the form of Alibaba, Badoo, and Tencent. But as Jeffrey Ding has pointed out, China is not monolithic. It is strong, but not dominant in key areas of AI innovation, like computers, chips, and algorithms.

When will it eat us?

Finally, a word about being eaten. Remember Putin's question about "When will it eat us?" Some AI philosophers divide AI into the three categories. There's the current state of affairs—narrow AI—that's when an AI application is generally better than a human at a particular task, like pattern analysis.

AGI, or artificial general intelligence, also known as HLMI, human level machine intelligence, represents a point in time—really a phase in time—where an artificially empowered machine is generally better than humans at a number of tasks, can move fluidly from task to task, and can train itself and write code. Then there's artificial super intelligence, which is reached when machines are smarter than humans across the board. A super intelligence machine can plug into the internet, consume the knowledge of the Internet, attach to the grid, and if necessary fool you into thinking it is safe, or so the debate goes. How does this end up with you getting eaten?

Enter the paperclip machine, a machine that neither hates nor likes, but simply is trained to make paperclips and optimized to do so. Eventually the machine runs out of grid-based energy, so it looks around for new sources of energy, and sees in a room like this a ready source of carbon energy, aka—humans. So it programs machines to collect the carbon so as to make ever more paperclips.

Let me be clear, if I were briefing a national security official in the United States government, I would not start or end with the paperclip machine. What concerns me about Putin's statement is that it means he's being briefed at the highest levels of the Russian government on some form of advanced AI. He is into the literature. So, why might Vladimir Putin get briefed on AI?

National security applications

Let's begin with some of the ***military applications*** for AI, some potentially transformative. Here, people tend to immediately think of robots.

The Russians are indeed making a robot they call FEDOR. It can shoot and carry heavy weights—like infantry—or if they are really heavy weights and good shots—Marines.

But there's more to it, including autonomous weapons systems and lethal autonomous weapons systems.

A couple of points here: 1) We have had autonomous weapons systems since the 1970s. To the military, at least, the concept of an autonomous weapons system is not new. 2) The Department of Defense (DoD) has gotten the memo about AI. In 2014 DoD initiated something called the Third Offset. An offset is DoD speak for consciously harnessing the nation's technology to offset an opponent's advantage.

The first offset took place in the 1950s and sought to compensate for Warsaw Pact advantages in manpower in Europe through the development and deployment of tactical nuclear weapons.

What is notable about the Third Offset is that it's a technological effort that doesn't seek to offset the opponent's advantage with a weapon, but through technology generally, and in particular AI.

If you want to imagine how AI might enable offensive and defensive weapons, think of swarms of AI controlled birds, objects, or bits of metal working in unison to attack an aircraft carrier, as Kamikazes once did, but in a coordinated manner and without the moral or supply issues involved. Likewise, swarms can serve as super chaff confusing and diverting incoming air and naval missiles.

What else can AI do in the military context? Logistics. You want to plan D-Day? Imagine planning the D-Day invasion, but doing so with a D-Day Waze App that tells you exactly what to load on which ship when to land and then makes instantaneous adjustments for weather and intelligence input.

You can use AI for war gaming. You can test weapons with it. In general, any military task that involves danger or repetition or that might diminish in quality in response to fear or fatigue may be improved with AI applications.

Now let's turn to intelligence. One of my favorite AI writers is Ryan Calo (University of Washington). He writes in English and understands the links between law, policy, and technology. He says the whole purpose of AI is to spot patterns people cannot see. Spotting anomalies and patterns is the business of intelligence as well.

Consider what AI can do for image recognition, voice recognition, sorting, aggregation, prediction, translation, anomaly identification, and everything that occurs in cyberspace.

Now imagine you are trying to figure out if North Korea is violating a sanctions regime, AI can aggregate all the information available electronically, about shipping routes, bills of lading, and so on, track them and find the anomalous pattern.

If you're doing social media analysis, AI can aggregate ISIS Facebook posts, and find patterns in messages and locations. AI will also shape watch lists and power biometric recognition and detection.

AI could also transform ***decision-making***. The hardest thing to do in government, other than to make the decision, is to first fuse the information that should inform the decision. In my day, the best way to fuse intelligence was to have a Principals meeting, and see what intelligence the Secretary of State, the Secretary of Defense, and the DNI brought to the table. That's not a great method. AI can do this better than any other method I've seen.

All that's great news if you're on our side of the table, perhaps. Now think about how AI can help an authoritarian regime maintain control, perhaps with something called a social credit system. You all know what I'm talking about, that's the new Chinese mechanism of social control.

If you jaywalk, you are called out. If you post negative comments, your score is lowered. But lest you think this is just a Chinese feature, in the UK, if you talk on your handheld cell phone while on the highway, you will receive a license suspension in the mail, as well as a photo of yourself on your phone.

With Fake App, imagine what you can do if you're destabilizing another country's democratic elections. Alternatively, I urge everyone in this room to read the February 16, 2018, 37-page Justice Department indictment of 13 Russian agents describing some of the media based methods Russian agents used to interfere in the 2016 election. Now multiply by an AI factor of 10 or 100.

Non-state actors may pursue AI as well—think about how AI can be used to enhance IEDs. The swarm concept could replace the cinderblock-on-gas-pedal method of delivery.

National security risks and implications

AI, Stephen Hawking noted, has many potential benefits, as well as risks. Already, for example, AI pattern recognition is better at detecting and distinguishing between benign and malignant tumors than humans. AI has the potential to address and solve disease and challenges of environmental degradation. But I am addressing AI from a national security perspective, so that means I am focused on risks, as well as benefits, and I tend toward worst-case scenario planning.

When it comes to national security decision-making, for example, I am thinking about how AI will compound existing decisional pathologies, like cognitive bias, speed, secrecy, and what I call the national security imperative, the drive to solve the national security problem at the risk of long-term consequences and values.

For sure, AI is going to mitigate some of the pressures that come from real world deadlines and decision-making with too little information and too little time. But it could also shorten the amount of time people have to make decisions, and in dramatic fashion.

Think about hand-to-hand combat in cyberspace, and if that's AI enabled, and there's no person in the loop making decisions.

Likewise, it is an advantage to have AI capacities as a probability predictor, but it will only work if the people using AI —and that means people like me, not people like you, who are less comfortable with technology—understand why it is that there's a 76% probability that someone is a terrorist, and that this is a prediction and not necessarily a fact.

Policymakers love to ask intelligence people, are you 87% sure? Are you 93% sure? And the intelligence people want to say, it is our informed judgment that…

I worry that AI could drive some decision-makers to look for mathematical answers. Fair enough, if it is used as an augmentation tool. But some national security making is intuitive, whether to go or not on D-Day. Plug that into an algorithm, and we'd still be waiting for perfect weather before making the cross-channel landing.

Military command & control implications

The US military currently uses what they call the Centaur model of AI command and control. That means AI is used to augment human capacity, not replace it with completely autonomous weapon systems. In that regard, Secretary of Defense Ash Carter enunciated a no-first-use policy with respect to AWS, that is, the US would not deploy fully autonomous weapons systems in the first instance, but would reserve the right to respond in kind. Doctrine itself could pose a risk, theirs and ours.

The law of unintended consequences may come into play as well. Technology does not always work as intended as Icarus learned the hard way. But you do not need mythology to make the point. You can look to the Mark 14 and 18 torpedoes, the former of which had a propensity to hit enemy ships and not explode, tending to give away one's presence and position.

Post-Sputnik, the US ran out and launched two satellites. Both exploded on the pad. The list continues with Apollo 13, the Challenger, and Columbia. Stuxnet worked as intended, but then it jumped the rail.

The reality that technology does not always work as intended underscores the importance of having a human in the loop. AI specialists are all familiar with the 1983 Petrov incident. Lieutenant Colonel Stanislav Petrov was the watch officer at a Soviet Rocket Force command and control center, when radar indicated the launch of a US first strike. Petrov was on the clock.

By doctrine, he was required to immediately inform his chain of command, up to the Politburo. He delayed; he did not think it looked like what a first strike would look like. He was under great pressure to act. It took a lot of guts, especially in the Soviet system, for a lieutenant colonel not to follow procedure. He did not, and soon a technical error was revealed.

Lest the incident be attributed to Soviet technology, in 1979 National Security Advisor Brzezinski experienced a similar episode. The DoD Command Center woke him in the middle of the night to the news of a Soviet weapons launch and began a countdown to inform the president and respond. Here too, the system corrected itself before Brzezinski reached the president.

Sydney Freedberg and Matt Johnson have illustrated a third risk –even when technology works as intended, there may be interface issues between man and machine. They use as example Air France 447, which crashed in the Atlantic going from Brazil to Paris, in part because of the inability of the pilots to transition from autopilot to manual flight in the midst of an emergency. The authors also cite a couple of friendly fire instances involving Patriot batteries in 2003. The technology worked as intended. But when the technology passed control back to the human in the loop to make a fire or no fire decision, the operators were not sure what was happening and made erroneous choices.

Likewise, the Vincennes incident, involving the shooting down of an Iranian commercial Airbus by an Aegis cruiser is an example where the technology worked, but the interface did not. Human actors misread the data. By the way, an AI-enabled system would likely have prevented the tragedy.

That AI can be programmed to instantaneously respond is both an advantage and a disadvantage.

Foreign relations impact

Some economists predict that a significant portion of the male populace of the world will be pushed out of work by AI. Larry Summers, for example, has stated that by 2050, as many as one third of the world's 25- to 54-year-old male population will be out of work. This is an obvious source of instability and national security risk.

AI may also exacerbate current North-South divides as well as create new ones. It can give new power to smaller states, a version of the Singapore effect. AI will create supply chain and counterintelligence issues, illustrated by today's story about Qualcomm and the CFIUS process. And, as noted earlier, it may magnify the impact of asymmetric threats, like interference in elections as well as potentially strengthening the social control exerted by authoritarian regimes.

Finally, there are risks associated with an AI arms race. There is a community of AI scholars that still hopes to avert an arms race. Indeed, one of the Asilomar AI Principles is that "an arms race in lethal autonomous weapons should be avoided." AI think tanks talk about an AI Baruch plan, a reference to the plan proposed by US UN Representative Bernard Baruch at the close of World War II that would have had the US and other states turn over their nuclear weapons capacity and know-how to United Nations control.

But there will not be an AI Baruch plan, because there is too much to be gained from AI, too much security advantage. And, who by the way, is going to trust Vladimir Putin, who thinks that whoever controls AI will control the world. That is the national security imperative. An arms race is here with all the risk that comes with speed, shortcuts, and miscalculations.

Existential risk

Finally, there is the potential existential risk suggested by Elon Musk, Stephen Hawking, and Nick Bostrum among others. Essentially, there are three camps.

Camp one is represented by James Barrat, who wrote the book, Man's Last Invention. It's our last invention because AI enabled machines eventually will achieve mastery over humans leading to humanity's extinction, refer back to the paperclip machine.

Camp Two is the fork-in-the-road camp. AI could be friendly, or it could be unfriendly. This is where Stephen Hawking was, and it appears to express the perspective of The Future of Humanity Institute at Oxford.

Then there is the third camp—the stay-calm-and-carry-on camp. We'll fork right, or left, when the time comes. This is everybody at MIT, probably, and it's the people who have optimism this is all going to work out because they are going to write the algorithm that will make it work out.

But lawyers will notice some lingering nuance among the optimists. This is the Stanford 100-year Study on AI.

"While the study panel does not consider it likely that near-term AI systems will autonomously choose to inflict harm on people, it will be possible for people to use AI-based systems for harmful, as well as helpful, purposes….Contrary to more fantastic predictions for AI in the popular prose, the study panel found no causes for concern that AI is an imminent threat to mankind."

And then back to Ryan Calo. "My own view is that AI does not present an existential threat to humanity,"—awesome—"at least not in anything like the foreseeable future."

My own view? I do not worry about the longer term existential threat. There are enough near and intermediate threats to contend with, and if we don't address these threats well, and soon, we really won't need to worry about the long term existential risks.

How should we respond?

I have a helpful 157-part program. But in the interest of time, I would like to make three points and then offer a framework with which to address AI going forward.

First, we should start making purposeful decisions. National security policy should not be made by Google. That is wrong as a matter of democracy, and it is wrong as to how the competing equities are addressed. Mind you, national security policy should not be made by the FBI either, which leads to my second point. Litigation is a lousy way to make policy, which is where we are headed. If we do not fill the AI policy void, we will end up with endless iterations of the 2015 FBI-Apple litigation over access to the San Bernardino shooter's cell phone, a little league game compared to the litigation that will occur with AI where the stakes are so much higher.

For sure, litigation can serve as a forcing mechanism. But litigation accents the voices and interests of a few—the litigating parties—not those of society at large. The government's litigation process is also different from its policy process. It accents the views of lawyers making arguments to win cases, rather than those of policymakers addressing competing equities. Recall that some of the loudest critics of the FBI's position in the Apple dispute came from within the intelligence community. It is called the adversarial process for a reason. If you want informed and sound policy, better to do it in moments of calm dispassion, rather than litigation. Litigation also often results in divided rather than national policy, as courts divide over outcomes.

One more thing, the FBI-Apple dispute reminds us that where national security is concerned the government will make novel uses of the law, in that case, the All Writs Act of 1789. I think it is safe to say that Congress was not contemplating AI or iPhones at the time.

Third, ethics is not enough. I read you one of the principal statements from the Asilomar Conference AI ethical code, about avoiding arms races. That is not going to do the trick. I know from the Model Rules of Professional Responsibility for lawyers, the Code defines the basement of conduct, not the ceiling. I would also urge that you ask the MIT professors in this room just how constrained they feel by the professional ethical codes that purportedly bind them.

What should a legal framework address

National Security Law serves three purposes. It provides essential values. It provides essential process. And it provides the substantive authority to act, as well as the left and right boundaries of action. I am going to give you an example of each, which is all we have time for, but the overarching point is that the values-process-authority rubric should guide how we look at AI issues going forward.

Let's start with values, the most important national security and legal values in this area derive from the Constitution. The values that will be debated, litigated, and tested are the values embedded in the First, Fourth, and Fifth Amendments addressing government conduct.

First Amendment. Why the First Amendment? Think about analyzing Facebook postings for Russian interference or efforts to validate the authenticity of political ads. The First Amendment addresses five rights: freedom of the press, speech, religion, assembly, and the right to petition the government. Every time the government, in law or practice, takes an action that can be construed as impeding, restricting, chilling, or favoring one voice or view over another there is space for First Amendment challenge.

Think as well about the AI researcher who is traveling to the United States to speak at MIT and can't get in because there is a travel restriction. Or consider, potential disputes over funding which, depending on how it is allocated or withheld can create First Amendment issues.

Fourth Amendment. The Fourth Amendment protects you from unreasonable searches and seizures by the government, and in many cases, but not all, requires the government to get a court warrant before it engages in a search or seizure.

The key word is unreasonable. The debate is about what is reasonable in context. Since 1979, the answer to that question has depended in part on what is known as the Third Party Doctrine. The doctrine posits that if you share information with a third party, you do not have a reasonable expectation in its privacy if that information is subsequently shared with the government. Some courts have compared the doctrine to the attorney-client privilege, which protects communications between client and counsel for the purpose of legal representation, unless the information is shared with a third-party in which case the privilege is said to be waived. But the Third Party Doctrine is different.

If I have a cell phone and I want to call someone, I necessarily have to share my calling data with the carrier in order to place the call. Under the Third Party Doctrine I have lost my reasonable expectation of privacy in that data vis-à-vis the government. Never mind that the case this all depends on dates to 1979. It involved the intimidation of the victim of a purse snatching. The government trapped and traced the number of the person calling the key witness to threaten her. The government offered the calling data into evidence without having obtained a warrant to seize it. The Supreme Court said fine, there was no reasonable expectation of privacy with information shared with the third party phone company. That's point one. We are relying on dated case law.

Point two, the government generally takes the view that if it lawfully comes upon your information, it may use it as it sees fit, or as otherwise permitted by law. Thus, barring some other law, public images gathered on the Internet, for example, might be used for AI machine learning. Likewise, Congress in 1986, passed the Stored Communications Act, which further regulates certain Third Party information shared with ISP's by requiring a warrant before the information is shared with the government.

The US Supreme Court is hearing a case right now called Carpenter, involving cell phone tower location data that could put parts, or all, of current doctrine in play, because the Supreme

Court has grown nervous about the aggregation of information, even if that data falls within the third party doctrine. Alas, the case is ill-suited for changing settled law.

Why does all of this matter to AI? Because AI is predicated on large data sets and most of the legal values expressed in the law today are found in the Bill of Rights, which regulates government access, but not private access and use, where much of the AI data is generated, stored, shared, and used. Moreover, where there is law, it is out of date. Finally, criminal litigation over Fourth Amendment searches is a poor forum in which to resolve national questions about AI data and privacy.

Fifth Amendment. There are two clauses with significant AI implications. First the Fifth Amendment provides that the government cannot deprive someone of life or liberty without due process of law. In very short order, that essentially means an opportunity to be heard and to make your case. Deep machine learning and AI implicates the clause, because AI algorithms may operate within a neural network where it is known what input went into the machine and what output came out, but not exactly what occurred between, as algorithms weighed and sorted data across neural networks. The question arises, to what extent does the person impacted, perhaps during a parole hearing, or in watch list litigation, have a due process right to look beyond the predictive result to the manner in which the result was reached. And, what if the judge and lawyers in question do not understand what occurred themselves. Will courts accept the results? Require expert testimony?

Second, the Amendment expressly provides that the government cannot take someone's property without paying just compensation. How might that come up in AI context? Stipulating that most of the AI research in the US is conducted by industry and academia, imagine that one such entity creates a process or algorithm of national security importance. The takings clause would seem to suggest that the government could seize the property, provided it paid just compensation for doing so. To state the obvious, there are values as well as valuation issues involved.

The second purpose of law is to provide essential process. You can't have process without structure and organization. Here, if I were to pass one law right now, or issue one directive, it wouldn't deal with substance, it would address how the US government is organized to address AI. The present organization is not manifest. I don't know who speaks for the government on AI. Moreover, the lead bill in Congress creates an advisory committee to advise on what the structure should be.

But the AI train has left the station and it is time to get moving and start making timely and purposeful decisions about: Who will take the lead? Do they have the necessary authority to speak for the government? Before Congress, in Silicon Valley, and at Federal Funded Research Centers? Do they have authority to resolve disputes between major departments and actors, or the wherewithal to readily raise these disputes with the President? I would like to skip the ten years of debate and get on with the decisions.

Finally, what of the substantive authority to act and right and left boundaries. There are too many laws to address in this short forum. Some are consciously addressed to AI, most are applicable by implication.

Here are a few of the laws that matter most in defining the limits and authorities of governmental authority: The Defense Production Act, The Inventions Secrecy Act, and The International Economic Emergency Powers Act.

There is also much to be learned from Nuclear, Chemical, Biological arms control concepts in domestic and international law as well as law of armed conflict concepts like command responsibility and new weapons review.

Beyond identifying the law, my purpose here is to encourage scholars and commentators to consider these laws in AI context in a way they have not done before.

I also want to make the obvious point that the law never keeps up with technology. Moore's law is always faster than Congress's law and case law. But if this is an obvious lesson, why do we keep learning it.

One answer is to focus on laws that impose process and checkpoints, like FISA, rather than seek to dictate substantive results, or impose specific permits and prohibitions addressed to specific technologies that may evolve—will evolve—in unanticipated ways with unanticipated uses.

Conclusion: Looking ahead

The title of this presentation is A Dangerous Nonchalance. It feels that way to me. The Belfer/IARPA Study states that AI is likely to be as transformative a national security technology as aviation or nuclear weapons. Some of the AI philosophers go a bit further and suggest an existential threat to humanity itself. And yet, I do not feel a sense of urgency to address the legal, ethical, and policy challenges ahead.

I can guarantee you it is not what people are studying in law school. I do not sense that it is what students are studying at War College either. During the Cold War, it was understood that a security specialist would understand nuclear doctrine. And, there was doctrine to understand. Time to do so now with AI.

In 2016 scholars at the Future of Humanity Institute, AI Impact, and Yale asked AI experts from around the globe representing academia, industry, and government when AI might exceed human performance, Human Level Machine Intelligence. The median answer for respondents from China was 28 years. The median response for American specialists was 76 years.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

STARR FORUM REPORT

MIT CENTER FOR INTERNATIONAL STUDIES
Building E40-400 | 1 Amherst Street, Cambridge, MA 02139 | cis.mit.edu

May 25, 2018
18-01

Artificial intelligence and national security law: A dangerous nonchalance

Honorable James E Baker, MIT

We need to talk and bridge the gaps now, government-to-industry-to-academia and between technology-policy-and-law. It is time for purposeful decision, and policies made through the democratic, rather than commercial or litigation process.

There are two overarching questions that go well beyond the role of government. First, if technology creates risk, how can AI technology be used to minimize risk? What are the core technological questions? What are the limits of AI coding? Is it possible to build in AI attribution? What are, or should be, the technological red-lines? And, how will you explain the answers to these questions in plain policy and legal English?

Finally, what is the responsibility and role of a US corporation? There is no common understanding, as there once was. And, there is likely no correct answer. But there are incorrect answers, like default answers, and answers that are derived not from frank and full debate and purposeful decision embedded in law and ethics.

Massachusetts Institute of Technology
Building E40-400
1 Amherst Street
Cambridge, MA 02142

MIT CENTER FOR INTERNATIONAL STUDIES

NON PROFIT ORG.
U.S. POSTAGE
PAID
Cambridge, MA
Permit No. 54016

MASSACHUSETTS INSTITUTE OF TECHNOLOGY