May 30, 2019

To:

AI-Standards, National Institute of Standards and Technology
100 Bureau Drive, Stop 2000, Gaithersburg, MD 20899.
E-mail: ai_standards@nist.gov

Re: RFI: Developing a Federal AI Standards Engagement Plan

This comment is in response to the Request for Information (RFI) about Federal Engagement in Artificial Intelligence Standards issued by The National Institute of Standards and Technology (NIST) as directed by The February 11, 2019, Executive Order on Maintaining American Leadership in Artificial Intelligence (AI).

MIT recently announced[1] they are committing $1 billion to address the rapid evolution of computing and AI and the associated global opportunities and challenges. However, AI is no longer constrained to academia. According to a report[2] by CB Insights, "Artificial intelligence is changing the fundamental structure of every industry in areas ranging from agriculture to cybersecurity to commerce to healthcare, and more." Furthermore, "Governments are competing to establish superior AI research, seeing AI as a lever for greater economic influence and power."

AI is such a potentially powerful technology that in a recent report[3] by the Federation of American Scientists, it was listed as one of seven emerging technologies that have significant potential for substantially dangerous or disruptive effects on strategic nuclear stability. The report emphasizes "artificial intelligence and its related technologies… as having potentially the greatest impact on strategic stability over the next 20 years." Beyond just strategic stability, "The global race to develop, own, finance, dominate, and disseminate artificial intelligence and other emerging technologies will permeate the business competition of the future and further divide nations based on their ability to capture these gains in a competitive global landscape."

While AI is defined broadly in the RFI, it may be characterized fundamentally as decisions made by a computer intended to simulate human behavior. Decisions, in this context, are logical paths that are defined by a human programmer. Each action performed by AI is a result of a preceding decision or series of decisions. Due to the ability of AI to "learn" through decision confidence improvement as a result of repeated processes, and the speed at which computers can make decisions and process data, it can be difficult to achieve clarity into AI systems. A lack of clarity ultimately hinders our ability to control AI systems. However, a robust set of AI technical standards is a means to help understand and control AI decisions. While standards outline the particular aspects of AI we wish to control, transparency is the foundation upon which standards must be built. Standards can only be imposed when AI decisions are well understood, and therefore, an effective set of AI standards should include a distinct set of transparency standards. Transparency[4] ensures that AI serves humankind and does not perform in a manner inconsistent with our goals. This means that while AI standards must be tied to underlying transparency, they must also be aimed at adherence to overarching frameworks that represent those goals, such as Moral and Ethical, Legal, and Safety and National Security frameworks. To ensure that AI

always serves the needs of humankind, and that the US can lead in AI innovation while protecting our values, a multi-tiered framework of AI guidelines and regulations should be established by the US government, with transparency as the foundation.

Federal involvement in AI standards should include a set of guidelines and regulations designed to inform and protect the public. For AI standards to promote AI innovation and US leadership in AI, and simultaneously protect civil liberties, privacy, American values, and United States economic and national security, an AI regulatory framework should be multi-tiered. A three-tiered regulatory system could consist of the following components depending on the AI application:

1. Advisory Only – For non-critical applications, defined as applications of AI that do not pose a significant risk to the public, federal guidelines and advisory notices could be issued with the goal of providing information and guidance to groups or individuals developing or interacting with AI systems. The advisory role lays the groundwork for understanding the important issues and considerations related to AI, while avoiding potentially burdensome adherence to mandatory guidelines that are not necessary. For research, development and testing in both the public and private sectors, federal guidelines and advisory notices are a way to foster innovation by engaging early with groups and individuals working to advance AI while avoiding imposing regulations that might slow progress.
2. Voluntary Certifications – For non-critical applications of AI that are public facing in some way and where public trust in a brand, service or general application of AI is desirable, a program of voluntary federal certifications could be created. If there is a particular trait of AI that is considered desirable, it can be communicated with a high degree of trust through a federal certification which acts as a stamp of approval. For example, in the food industry it may be desirable to communicate that a food and its ingredients are organic. While there is nothing to prevent the food manufacturer from using organic ingredients and creating an organic food product, the manufacturer may only claim their food is "USDA Certified Organic" if they follow the certification process set forth by the USDA[5]. The USDA stamp of approval provides the consumer with confidence in the claim that the food is organic. In a similar way, a federal certification would provide anyone interacting with an AI system a high degree of confidence that the AI is operating as intended.
3. Regulations – For critical applications, defined as applications of AI that could pose a significant risk to the public, regulations should be imposed to protect those potentially affected. This would include a federal approval process for public facing, critical AI systems. Examples would include AI systems used in autonomous vehicles, autonomous or semi-autonomous weapons systems, and any applications where bias could substantially disadvantage certain groups or individuals. For example, in the pharmaceutical industry, the FDA is responsible for regulating and approving drugs that are developed for public consumption[6]. The FDA's approval process is designed to protect consumers. Similarly, federal AI regulations would protect the public from potentially dangerous or flawed AI systems.

Federal guidelines and regulations would ultimately serve to ensure that the AI standards built on transparency and adherence to our goals are followed, and that innovation in AI is promoted and the public is protected. Establishing a robust set of guidelines and regulations may require committees and

working groups of experts to address key elements of a coherent plan. Examples of potential committees or working groups, which may overlap or be combined, include:

1. Committee on transparency – define transparency, may differ based on application (military versus private technology sector, for example).
2. Committee on standards – collect from groups existing standards that have already been developed, identify needs, work towards defining a complete set of technical standards.
3. Committee on adherence of standards to high-level goals – ensure adherence of standards to moral and ethical framework (examples include valuing human life, avoiding deception, avoiding unfair treatment of groups or individuals), US legal framework, and safety and national security framework. Work with independent groups, private enterprise, academic institutions and governmental bodies on establishing and defining frameworks to which the standards should adhere.

A robust and effective multi-tiered framework of AI guidelines and regulations will ensure that AI technical standards are well defined, built on transparency, adhered to and align with our goals and values.

Sincerely,

Jimmy Saldana, Director of Strategic Innovation
Arrowhead General Insurance Agency, Inc.
701 B Street, Suite 2100
San Diego, CA 92101
(619) 881-8595

Supporting References:

1. MIT reshapes itself to shape the future

http://news.mit.edu/2018/mit-reshapes-itself-stephen-schwarzman-college-of-computing-1015

2. Top AI Trends to Watch In 2018

https://www.cbinsights.com/research/report/artificial-intelligence-trends-2018/

3. Emerging Disruptive Technologies and Their Potential Threat to Strategic Stability and National Security

https://fas.org/wp-content/uploads/media/FAS-Emerging-Technologies-Report.pdf

4. Google says it will address AI, machine learning model bias with technology called TCAV

https://www.zdnet.com/article/google-says-it-will-address-ai-machine-learning-model-bias-with-technology-called-tcav/

5. Organic Regulations

https://www.ams.usda.gov/rules-regulations/organic

6. FDA Mission

https://www.fda.gov/about-fda/what-we-do