STATEMENT BY

DR. LISA PORTER
DEPUTY UNDER SECRETARY OF DEFENSE FOR RESEARCH AND ENGINEERING

BEFORE THE
HOUSE ARMED SERVICES COMMITTEE
SUBCOMMITTEE ON
EMERGING THREATS AND CAPABILITIES

ON

"Department Of Defense's Artificial Intelligence Structures, Investments, And Applications"

December 11, 2018

NOT FOR PUBLICATION UNTIL
RELEASED BY THE HOUSE ARMED SERVICES COMMITTEE

Chairwoman Stefanik, Ranking Member Langevin, and distinguished Members of the Subcommittee – thank you for inviting me to appear before you today to discuss artificial intelligence, particularly as it relates to national security applications.

As this Subcommittee knows, artificial intelligence, or "AI," is not a new thing – as long as there have been computers, there have been engineers who have dreamed about enabling machines to think the way humans do. In fact, the Defense Advanced Research Program Agency (DARPA) funded much of the early work in AI decades ago. These efforts led to "expert systems," such as tax preparation software, that people take for granted today. But these early systems were limited to very narrow applications and could not generalize.

Today we are experiencing an explosion of interest in a sub-field of AI called "machine learning," where algorithms have become remarkably good at classification and prediction tasks when they can be trained on very large amounts of data. There are numerous examples of successful applications of machine learning techniques; some of the obvious ones include facial recognition in photographs and voice recognition on smart phones. However, there has also been a significant amount of hype and confusion about the current state of the art. It is the USD(R&E) position that we must not abandon the tenets of scientific rigor and discipline as we pursue the opportunities that AI presents.

Today's AI capabilities offer potential solutions to many defense-specific problems; examples include object identification in drone video or satellite imagery, and detection of cyber threats on networks. However, performance must be assessed rigorously against quantitative metrics that are directly tied to the specific mission problem. For example, most commercial search applications focus on precision – meaning that, if I ask for images of cats, every image that comes back to me has a cat. Such algorithms may not do as well with recall – in other words, I may not get all the images that have cats in them. A metric that emphasizes precision over recall may not be appropriate for a military application where I am looking for, say, missile launchers, and I may be willing to accept some false alarms (lower precision) as long as I do not miss any launchers (higher recall). If I do not optimize my algorithm against the proper metric, I will not get the performance I need.

A related challenge to proper metric selection is determining the performance level required for operational utility. Oftentimes, current systems and capabilities are not quantitatively benchmarked, making it difficult to know what level of performance to target, and therefore how to assess whether the expected outcome will justify the development and system integration expenditures.

One of the drawbacks of today's machine learning techniques is the amount of human labor required to properly prepare, or "curate," the training data. For example, the impressive advances

in object detection and classification in imagery during the past few years came about largely because of the arduous two-year data labeling campaign led by Dr. Fei-Fei Li, which employed over 50,000 workers ("Mechanical Turks") from 167 different countries to generate about 15 million curated images. But while the performance of algorithms trained on this extensive data set against similar imagery is quite impressive, their performance against different kinds of imagery, such as satellite imagery, is not. In other words, despite all of the successful examples of current machine learning systems, they are narrow in what they can do, they are brittle, and they cannot explain what they do – which makes them hard to trust. Furthermore, current systems require robust processing power. And finally, current systems are susceptible to various forms of spoofing, known as "adversarial AI."

We are working to address these challenges and vulnerabilities – metric selection, data curation, trust, processing power, and adversarial AI – through multiple efforts, most of which will leverage the complementary roles of the Joint Artificial Intelligence Center (JAIC) and the USD(R&E) enterprise. The JAIC will offer a means to rapidly determine the appropriate metrics for operational impact for a variety of applications, and these insights will help inform algorithm and system development across multiple USD(R&E) research efforts. Furthermore, the JAIC's focus on scaling and integration will drive innovation in data curation techniques. A specific example of the synergy that we plan to foster between our organizations is a recent partnership between the Defense Innovation Unit (DIU) and the JAIC focused on predictive maintenance, where DIU chose a successful commercial airline industry supplier to prototype a six-month pilot program for E-3 Sentry aircraft maintenance. DIU and the JAIC are now working together to scale this solution across multiple aircraft platforms, as well as the Army's Bradley Fighting Vehicle.

In order to address AI's "trust issue," DARPA's Explainable AI program aims to create machine learning techniques that produce more explainable models while maintaining a high level of performance, while USD(R&E), together with the Service Labs and our international partners, is pursuing methods, tools, and techniques to enable rapid verification, evaluation, and certification of autonomous and AI-based systems. The High Performance Computing Modernization Program is designing new systems that will provide ample processing power for AI and machine learning applications on the battlefield. Finally, countering adversarial AI is one of the key focus areas of DARPA's $2 billion AI Next campaign.

Ultimately, as we look to the future, we anticipate a focus on developing AI systems that have the ability to reason as humans do, at least to some extent. Such a capability would greatly amplify the utility of AI, enabling AI systems to become true partners with their human counterparts in problem solving. As an example, an AI system with sense-making capabilities could advise a warfighter in a time-sensitive situation on what action to pursue, enhancing the decision-making process by discounting the human's own biases. This goal is the quintessential

“DARPA-hard” problem, and we anticipate many false starts as we pursue it over the coming years. Nonetheless, it is important that we continue to pursue this cutting-edge research, given the significant investments our adversaries are making in AI. We are therefore grateful for the leadership and support that the Members of this Subcommittee have shown regarding AI. We also appreciate the establishment of the National Security Commission on AI, whose charter is appropriately focused on key areas that must be assessed objectively to ensure that the US maintains a leadership position in AI-enabled technologies and systems.

Thank you for your interest in this important topic, and I look forward to answering your questions.