Reply Delete Junk Block ···

RE: A PLAN FOR FEDERAL ENGAGEMENT IN AI STANDARDS -DRAFT FOR PUBLIC REVIEW 2-JUL-2019

ZS

Zach Shaw <zach@calypsoai.com>
Fri 7/19/2019 3:59 PM
ai_standards

Calypso AI is a newly established company that provides managed services and software solutions to solve the important technical challenge of building trust in artificial intelligence systems. We wish to submit the following views on standards and tools development for AI systems that are in Appendix 5, which were part of the NIST Request for Information (RFI) on artificial intelligence standards (Docket Number: 190312229-9229-01), recognizing that the deadline for submitting comments on that RFI have closed. These comments are focused on the first eight topics of that RFI under "*AI Technical Standards and Related Tools Development: Status and Plans,*" starting on page 33, line 1029.

1. AI technical standards and tools that have been developed, and the developing organization, including the aspects of AI these standards and tools address, and whether they address sector-specific needs or are cross-sector in nature (Page 33, Line 1030-1032)

Calypso AI has developed a suite of tools that enables organizations, users, and other stakeholders to trust their AI, specifically addressing accountability, explainability, robustness, adversarial robustness, and reporting. This suite is cross-sector in nature - it addresses the needs of all AI systems, but the implementation of the tools is custom to specific sectors due to the varying needs for each sector for each portion of trustworthiness. Calypso AI's products vary by data type, which most significant experience in computer vision and cybersecurity-based data.

Calypso AI has also developed certifications for explainability and security which validate the need for internal understanding of AI systems and security of AI systems. That understanding enables improved performance of the system, resolution of security issues, and reporting to all stakeholders. The security certification ensures AI models are hardened against all types of adversarial attacks - evasion, poisoning, and model or data theft.

IBM has consistently released relevant standards-oriented research and open-source repositories. IBM released the following open-source validation efforts for adversarial robustness (CROWN, ART, CLEVER scoring) that attempt to generalize the robustness of AI systems, agnostic to the model type. IBM has also released a best practices-inspired framework for accountability of AI systems - the idea of including FactSheets which inspire trust by describing the lineage of a product along with the safety and performance testing it has undergone.

There are numerous AI international and national standards and working groups, both cross-sector and sector-specific - most notably the ISO and IEEE. NIST is already aware of many of these. Calypso AI encourages NIST to continue to use existing standards as a foundation for its standards. These standards will serve as effective guides for NIST's facilitation and development of standards, primarily towards asking the right questions.

2. Reliable sources of information about the availability and use of AI technical standards and tools (Page 33, Line 1033-1034)

Calypso AI is a reliable source about the use of AI technical standards in the space. In addition to our clientele, we have conducted extensive market research into the use of AI technical