June 10, 2019

AI-Standards
National Institute of Standards and Technology
100 Bureau Drive, Stop 2000
Gaithersburg MD 20899

Re: RFI: Developing a Federal Artificial Intelligence Standards Engagement Plan
Response to 84 Fed. Reg. 18490 Request for Information dated May 1, 2019

Dear Sirs/Madams:

The following is in response to the request by the National Institute of Standards and Technology ("NIST") for industry input to help create a plan for Federal engagement in the development of technical standards and related tools in support of reliable, robust, and trustworthy systems that use Artificial Intelligence ("AI") technologies. Specifically, NIST has requested information to understand the current state, plans, challenges, and opportunities regarding the development and availability of AI technical standards and related tools, as well as priority areas for federal involvement in AI standards-related activities.

It is important to note at the outset that there are a number of products and services that comprise the value chain for AI that should be included as part of this review. AI technologies – technologies that enable machine-led analysis of large amounts of data to produce 'intelligent' results – are by definition dependent upon the appropriate collection, storage, organization and management of large amounts of data. It is these underlying data-related technologies that form the necessary foundation for all AI technologies.

As a leading private sector provider of these data-related products and services, NetApp is in a unique position to provide requested input in this process. NetApp has over 25 years' experience helping customers with solutions to help store, protect, replicate and secure data. In addition, NetApp has played a major role in advancing industry standards that are ubiquitous today, including but not limited to Network File System (NFS), Network Data Management Protocol (NDMP), Internet Small Computer System Interface (iSCSI) and Fibre Channel over Ethernet (FCoE). NetApp continues to drive new and emerging standards in cybersecurity, solid-state storage, non-volatile memory, power efficiency, virtualization, cloud computing, high-performance computing, and open-source initiatives and other technologies, many of which enable AI technologies. Accordingly, we are pleased to provide our input to support this Request for Information.

AI Technical Standards and Related Tools Development

We believe that there is a need for AI technical standards and related tools to support important goals such as interoperability and efficiency in the development of AI technologies, and want to highlight the interdependence of these technologies with other products and services in the overall value chain (e.g., data storage, management and organization). Given that the private sector

drives innovation in this area, we support and are appreciative of NIST's public, transparent and collaborative efforts with the private sector to develop such standards and tools.

Successful execution and use of Artificial Intelligence technology depend on data. Therefore, as noted above, there are a number of standards relating to data storage, management and organization that are critical to enable AI technology and are appropriately included within this review. We highlight a number of these and other areas of recommended industry best practice below for consideration and review in this process. All of the characteristics and best practices described below would be applicable regardless of industry sector.

- **Utilizing open source.** Much of the data used in AI technologies is unstructured, including both object and file data. As a result, we recommend that NIST include within its review, and encourage the use of, open standards related to file system data storage and object data storage such as Network File System ("NFS"). Because NFS is an open standard, it has been implemented by a wide selection of platform providers and is also deployed in multiple IT environments – e.g., on appliances, in enterprise data centers and in cloud services. Best practice for the use of AI would be NFS implementations that span IoT/edge processing, core datacenter processing as well as native cloud services

- **Moving data close to the computational power required for parts of AI technologies.** The ability to examine raw data in detail and by location is critical for AI applications. Raw data to be used by AI technologies must be quickly accessible and available with very high performance in order to meet the speed demands of today's Graphical Processing Units (GPUs). Best practice to meet AI's location data management challenges would be an architecture like a *Data Fabric* that provides capabilities of fast performance, capacity, and data availability across all locations in an AI lifecycle and which would include on-premises datastores, hybrid, and cloud facilities. Data Fabric architectures allow easy data movement across all locations to ensure optimal data locality and AI data processing.[1]

- **Interoperability capabilities to address data availability across multiple locations and platforms.** Because data may be resident in different locations throughout the AI lifecycle, best practice would include capabilities that are OS-agnostic and support all storage protocols to enable the broadest range of solutions. Such industry-leading requirements would include multiprotocol capabilities allowing data access across multiple host operating systems when needed, including all Linux and Windows operating systems to include file and block storage (SMB, NFS, iSCSI, FC, NVMe/FC).

- **Optimization of IO speed and raw performance**. AI technologies are naturally dependent upon the infrastructure upon which they ride, and it is essential in order to meet the speed demands of today's GPUs to maintain very high processing speed and

[1] Technical White Paper "Edge to Core to Cloud Architecture for AI Key Considerations for the Development of Deep Learning Environments" by Sundar Ranganathan, NetApp August 2018 | WP-7271, https://www.netapp.com/us/media/wp-7271.pdf.

performance. Though optimum speed and level of performance may vary by use case, best practice would be to ensure the underlying infrastructure provides optimal speed and performance capabilities.

- **Ensuring data quality by repeated testing and ensuring appropriate copies and freshness of valid data**. A Forrester study on impediments to AI technologies indicated that data quality is a key issue (in addition to talent and trust in AI systems).[2] Best practice for AI architectures would therefore provide the ability to rapidly and efficiently maintain data copies to ensure data quality. NetApp addresses this with capabilities that have become a de facto standard of enterprise data storage environments including the ability to snapshot and clone (point in time) copies of data.[3] In addition, the ability to maintain the freshness of data and data change rates of the AI data baseline is also important. For optimal efficacy, AI architecture must support the ability to quickly load, access, update, and refresh data to provide timely and accurate results. Finally, best practice would be for instant data 'snapshot' and clone backup and recovery capabilities to be widely available across all data locations in an AI data pipeline – on-premises, hybrid cloud and cloud facilities - to allow instantaneous data processing and re-processing for more and faster AI outcomes.

- **The ability to monitor, update and expand systems after deployment**. Flexibility is another critical requirement to support AI technologies, and therefore purpose-built solutions would not be recommended. For optimal efficacy, best practice would be for AI technologies to be able to update, expand, and grow the overall system, data sets, and capabilities without disruption while maintaining common data management practices.

- **Ensuring appropriate and reliable security.** Data security is a cornerstone of the Federal Data strategy and recommended action plan.[4] Therefore, best practice would be for AI architectures and platforms to support the ability to secure data through encryption and proper key management across an AI data pipeline, and for these security features to be built in to platforms rather than requiring add-on products and services. In some cases, certain industry sectors, such as regulated industries like healthcare or the intelligence community, will have specific or heightened data security requirements that will need to be met in addition to standard data encryption. Applicable standards in this regard would include FIPS and NIS compliant AES-256 data encryption as well as KMIP

---

[2] Forrester InfoGraphic, "AI Experiences A Reality Check Three Challenges Hold Firms Back From Achieving Enterprise AI Aspirations" by Michele Goetz and Elizabeth Cullen with Gene Leganza, Katie Hampton, and Chandler Hennig May 17, 2019, https://www.forrester.com/report/Forrester+ Infographic+AI+ Experiences+A+Reality+Check/-/E-RES153100.

[3] "If Your Data is Bad, Your Machine Learning Tools are Useless" by Thomas Redman, April 2, 2018, Harvard Business Review, https://hbr.org/2018/04/if-your-data-is-bad-your-machine-learning-tools-are-useless.

[4] "Draft 2019-2020, Federal Data Strategy Action Plan," https://strategy.data.gov/assets/docs/draft-2019-2020-federal-data-strategy-action-plan.pdf.

compliant key management capabilities, as utilized in NetApp's ONTAP data management platform.

- **Ensuring appropriate access, policies and data governance polystores.** The concept of data lakes or data oceans are commonly envisioned and deployed to store and manage data used for training and inferencing in AI workloads. Standards must be developed that allow data to be stored natively in SQL, NoSQL, File-based and Object storage methodologies. These dissimilar storage modalities represent a challenge in implementing standardized access control, information protection and data anonymization. Data lakes may contain PII, HIPAA, credit, transaction history, financial records, surveillance (video, photo audio) and U.S. Government classification markings. Each data type is unique and governed by standards for access, information protection and information lifecycle management. Research efforts into building and defining polystores represent an opportunity to leverage the power and growth of commercial storage solutions by overlaying a semantic notion of a unified data model and query capability. Conceptually standards can be devised at the polystore unification layer to provide granular access control, implement data governance, enforce policy and protect personal rights. Logically to the AI computation engines, the islands of dissimilar information represent a set of standards-based information engines which are accessed with a single query language simplifying the data preparation, transformation and normalization efforts that comprise the bulk AI development and deployment.

- **Pre-tested and validated infrastructure.** In order to optimize AI deployment in the Federal Government in particular, including providing agencies with the ability to focus on the data and outcomes versus the infrastructure, minimize deployment risks, and accelerate delivery/implementation time, best practice would optimally be for AI technical architecture guidelines or standards to be pre-tested and validated.[5]

Figure 1 provides an illustrated sample of an AI Deep Learning Data Pipeline and demonstrates why the design attributes described above are critical to optimal deployment of AI. As indicated, each "phase" has specific characteristics. Best practice would include (1) enabling deployments that span on-premises and cloud services, and not be limited to one or the other; (2) optimizing the management of data flow across these environments; (3) enabling data availability across multiple locations and platforms in order to allow varied applications and systems to consistently process data and achieve better results; and (4) utilizing platforms and architectures that offer multiprotocol capabilities allowing data access across multiple host operating systems, file and block storage (SMB, NFS, iSCSI, FC, NVMe/FC) and cloud services.

[5] Example: NetApp Verified Architecture ONTAP AI – NVIDIA DGX-2 POD with NetApp AFF A800 NVA Design David Arnette and Sung-Han Lin, NetApp May 2019 | NVA-1135-DESIGN | Version 1.0, https://www.netapp.com/us/media/nva-1135-design.pdf.

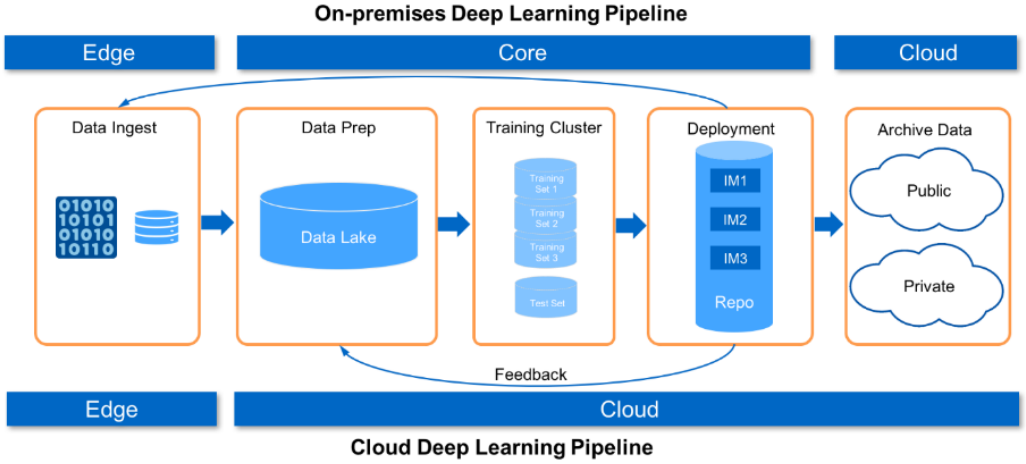

Figure 1 - Sample AI/Deep Learning Data Pipeline

Finally, it is worth emphasizing that the solution to challenges in the AI arena will be achieved by a combination of people, processes, and technology. While codifying data management tools and approaches is a critical step, it is also vitally important that practitioners in the AI field first take the time to closely study the problem they are trying to solve then see if it is possible to generate or capture the appropriate data that will allow them to produce reasonable, robust, repeatable, consistent, secure outcomes with the proper metrics to evaluate results. This is difficult and, in some cases, an effective AI solution may not yet be achievable.[6]

* * * * * *

We very much appreciate the opportunity to provide input to NIST as it helps to create a plan for Federal engagement in the development of technical standards and related tools in support of reliable, robust, and trustworthy systems that use Artificial Intelligence technologies. We are pleased to share our expertise throughout this process to assist the agency as may be needed, including providing any additional detail and/or use cases for the best practices outlined above, and look forward to participating in any future working groups, workshops or other venues provided for private sector input. Please do not hesitate to contact me directly with any questions regarding the foregoing.

Sincerely,

Kristen Verderame
Government Relations

---

[6] Derived from comments by Lisa Porter, Deputy Under Secretary of Defense for Research and Engineering at GeoInt 2019. *See also* https://www.meritalk.com/articles/ai-still-far-away-from-mission-critical-role-dods-porter-says/.