NVIDIA appreciates the opportunity to submit comments in support of NIST's mandate to develop a plan for the federal government's involvement with AI standards-related activities.

At NVIDIA, we engage with customers across a variety of industries, such as cloud service providers, healthcare providers, and automotive companies, as well as across many departments in the Federal government. This broad engagement gives NVIDIA a view to different use cases, approaches, needs and levels of AI expertise and adoption in the community.

NVIDIA has deep technical expertise in many areas of AI -- NVIDIA technologies have been credited with being a key enabler of the current AI renaissance. In particular, NVIDIA's expertise with AI Infrastructure, the hardware and software systems that enable reliable, scalable and traceable AI system development and deployment, prompts us to make several points:

- The development of AI systems at scale is quite different than traditional software development. AI system development is data centric as opposed to code centric, highly iterative in nature, and works best at scale with fewer humans in the loop.

- The government cannot simply "cut and paste" existing AI tools from industry and the open-source community for its own complete solution. This is due in part to the variety of data modalities, as well as security and traceability requirements across the government.

- NIST can work to discourage the replication of effort and the proliferation of silos of processing and data formats that will undoubtedly take hold as different organizations begin to adopt AI at scale for production use.

- NIST can work to ensure that data and AI models developed across different departments and parts of the ecosystem will indeed be interoperable, and that key learnings are encoded into AI models and not lost when there is a shift change or personnel rotation.

NIST could further these goals by helping to develop and promote AI *reference architectures*. These architectures would be guidelines for creating systems for managing, developing and deploying AI systems at scale. They would ensure that those who followed the reference architecture would develop models and datasets that were usable by others, and were following best practices regarding AI dataset curation, model validation, verification, and more.

At the same time, a reference architecture is not overly prescriptive regarding the underlying technologies. For example, AI processing systems should be deployable anywhere, in public, private or hybrid instances of cloud computing environments.

NVIDIA has significant experience supporting the data processing necessary to train and validate industry-grade AI systems at scale. As the government is beginning to transition its AI efforts from R&D to large scale deployments, there is both a challenge and an opportunity around interoperability and deployment of best practices.

We look forward to continuing this conversation.