# NASA STTR 2017 Phase I Solicitation

## T11.01 Machine Learning and Data Mining for Autonomy, Health Management, and Science

**Lead Center: ARC**

**Technology Area: TA11 Modeling, Simulation, Information Technology and Processing**

Nearly all engineered systems in all of NASA's areas of interest have one key aspect in common---they generate substantial data. These data represent:

- Science and scientific applications.
- The operations of the data collecting instruments and their platforms.
- The health of these instruments and platforms.
- In some cases, other related data such as the performance and health of the humans involved in operations.

Machine learning, data mining, big data, and related methods have been used to study data in these four areas individually for offline study, with the goal of understanding how the system really operates, as distinct from how it was designed and intended to operate. However, these data-driven methods have not been used so far to study data across more than one of these four areas, and not during operations, with the goal of enabling a human and/or autonomous system to make adjustments to the system's operations on the fly. Allowing both online and offline learning would allow for both online (tactical) and offline (strategic) adjustments to operations. Allowing humans and autonomous systems to interact in making strategic and tactical decisions, including user interfaces that allow the autonomous system to show the human what it has learned and the human to specify high-level objectives and/or low-level actions, is a key problem to be addressed. Increasing the scope of the data covered to all of the four areas above would allow autonomous systems and human operators to account for both science and system health drivers in operations, and identify the trade-offs between increasing science operations, increasing availability, maintaining systems health, minimizing maintenance costs, and other considerations. Some of these considerations may extend to improvements in on-demand system responsiveness through optimal resource sharing of the computational burden between online and offline computing platforms. Integration of learning autonomous systems into existing mission operations and systems is a key problem that will need to be addressed.

The utilization of the above types of data to optimize all aspects of operations is important for missions/projects in all of NASA's areas of interest such as space science (e.g., Kepler, TESS), space exploration (human and autonomous rovers), Earth science (satellite-based and airborne instruments and platforms), and aeronautics (e.g., UAS in the NAS) to operate them in as cost-effective a manner as possible. This becomes more critical as NASA increasingly moves towards operating multiple platforms in a coordinated manner (e.g., Distributed Spacecraft Missions, airborne Earth science platforms coordinating with satellite instrument platforms) where the volume of relevant data will increase and autonomy will be needed to properly operate the multiple platforms.

This subtopic has three goals:

- Increase the scope of machine learning, data mining, and big data methods within NASA to encompass both online and offline learning.
- Use data across as many of the above four areas of data as possible.
- Explore the trade-offs in operational efficiency, energy efficiency, health management, and operational performance/goal achievement between onboard and offboard computational resource platforms.

Proposed solutions may have characteristics including but not limited to:

- Ability to incorporate human feedback into the learning algorithms.
- Ability for machine learning algorithms to generate results for direct use by autonomous systems and human operators.
- Ability to learn a controller (covering strategic and tactical operations) from data representing human expert operations.
- Demonstration of a core set of tools that works across different domains.