

JUNE 2018

ASSURING OPERATIONS OF AUTONOMOUS SYSTEMS

**RON BIRK, ELIZABETH HORTON, THOMAS KASHANGAKI, ZIGMOND LESZCZYNSKI,
JON NEFF, NICK PERLONGO, TORREY RADCLIFFE, CHRIS TSCHAN, JOSH TRAIN
THE AEROSPACE CORPORATION**



ABOUT THE AUTHORS

All authors currently work at The Aerospace Corporation.

Ron Birk serves as Associate Principal Director for the Development Directorate in the Civil Systems Group and is a board member of the American Astronautical Society.

Elizabeth Horton serves as an Intellectual Property Specialist for the Civil Systems Group and is a member of the American Bar Association.

Dr. Thomas Kashangaki serves as the Systems Director for System Integration and Protection in the Space Science, Technology and Engineering Directorate of the Civil Systems Group.

Zigmund Leszczynski serves as the Systems Director for Technical Assessments in the Strategic Assessments and Studies Directorate of the Civil Systems Group.

Dr. Jon Neff serves as Systems Engineering and Data Science Specialist in the Strategic Assessments and Studies Directorate in the Civil Systems Group.

Dr. Nick Perlongo serves as Senior Technical Staff in the Strategic Assessments and Studies Directorate in the Civil Systems Group.

Dr. Torrey Radcliffe serves as Principal Director for Technology in the Civil Systems Group.

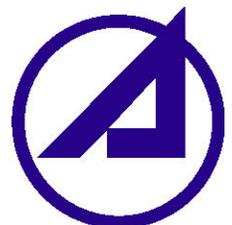
Dr. Chris Tschan serves as an Intelligent Systems Specialist, adapting intelligent systems techniques for forensic analysis of over 100 years of satellite state-of-health telemetry for 60 satellites.

Dr. Josh Train serves as the Principal Director in the Computer Technology and Research Subdivision in the Information Systems and Cyber Division of the Engineering and Technology Group, advancing and applying technologies to develop and acquire information systems.

ABOUT THE CENTER FOR SPACE POLICY AND STRATEGY

The Center for Space Policy and Strategy is dedicated to shaping the future by providing nonpartisan research and strategic analysis to decisionmakers. The Center is part of The Aerospace Corporation, a nonprofit organization that advises the government on complex space enterprise and systems engineering problems.

Contact us at www.aerospace.org/policy or policy@aero.org



Summary

U.S. aerospace agencies and companies employ complex systems-of-systems comprised of hardware, software, networks, and human-machine interfaces, with an increasing use of intelligent agents, artificial intelligence, and machine learning. These systems are evolving as “intelligent ecosystems”^a to operate autonomously, including the capacity to routinely upgrade themselves without human intervention. Assuring mission success requires system performance assessment fast enough to identify anomalies and take remedial actions to assure sustained and reliable operations. This paper informs decisionmakers on the urgency of establishing governance policies for verification and validation of system state-of-health needed to assure safe operations of autonomous aerospace systems affecting lives and property.

Domain of Autonomous Systems

Representative autonomous, intelligent ecosystems for civil applications include unmanned aerial vehicles, connected autonomous vehicles, space habitats deployed to the moon and Mars, space traffic management systems, environmental intelligence systems, and operations of complex nuclear facilities. Civil aerospace applications are leveraging developments and investments in autonomous systems as represented by the NASA Lunar Orbital Platform-Gateway shown in Figure 1. And from a defense perspective, according to Lt. Gen. Jack Shanahan, U.S. Air Force director of defense intelligence for warfighter support, intelligent systems with artificial intelligence are set to revolutionize how the military runs surveillance missions around the world.¹

^aAn intelligent ecosystem is a distributed, adaptive, scalable, system of systems with properties of self-organization, self-sustainment, and self-evolution. Definition provided by this paper’s authors.

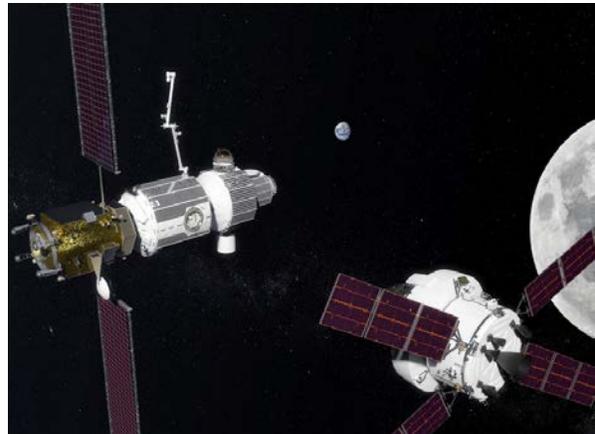


Figure 1. NASA Lunar Orbital Platform-Gateway (Source: NASA).

As configurations of space and other complex systems are programmed and networked to operate on a continuing basis, humans just cannot keep up with operating the systems in realtime. There is an immense amount of data flowing to monitor the surrounding environment, send commands, and make continuous adjustments for evolving operational conditions. Autonomous vehicles

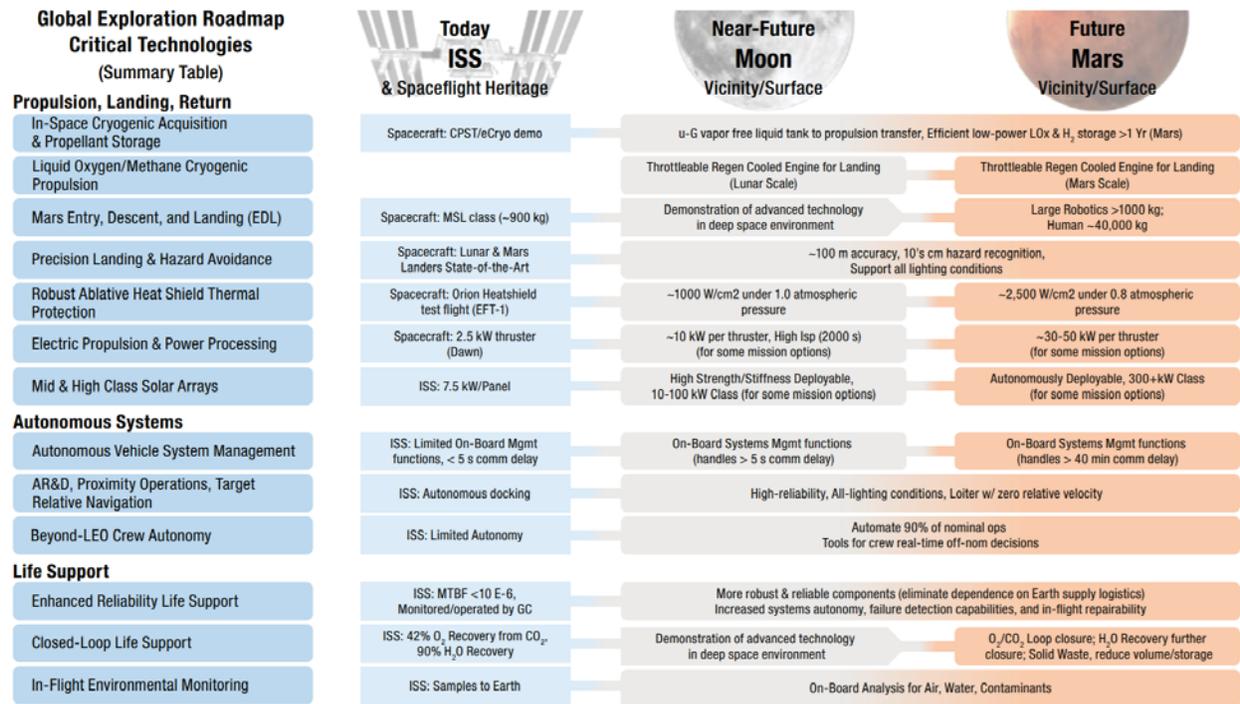


Figure 2. Segment of GER critical technologies roadmap highlighting Autonomous Systems (Source: ISECG).

maneuver using GPS, on-board sensors, and peer communications without needing human intervention. These intelligent ecosystems pass data, information, commands, and physical actions, and evolve over time through machine learning and artificial intelligence. Along with human-induced updates to software, hardware, networks, and intelligent agents,^b there is a need to track a system’s inherent evolution. System managers are challenged to maintain the current and accurate configuration to ensure safe, reliable performance and operations of autonomous, intelligent ecosystems over time.

As noted in the recently released Global Exploration Roadmap² (GER), “Autonomous systems enable the crew to conduct operations under nominal and off-nominal conditions independent of assistance from

^bAn intelligent agent (IA) is an autonomous system entity that directs activity using actuators in response to observations from sensors based on programmed goals.

Earth-based support. Advances in electronics, computing architectures and software that enable autonomous systems to interact with humans are needed and can be leveraged from commercial markets to support maturation of needed capabilities.” The GER contains a summary table of critical technologies (see Figure 2), including a segment of the table-highlighting autonomous systems.

Societal Impacts

As computing capacity, software performance, the Internet of Things (IoT), artificial intelligence, machine learning, network bandwidth, and robotic hardware continue to increase in capacity and connectivity, comprehensive and current knowledge of system state-of-health and susceptibility to threat

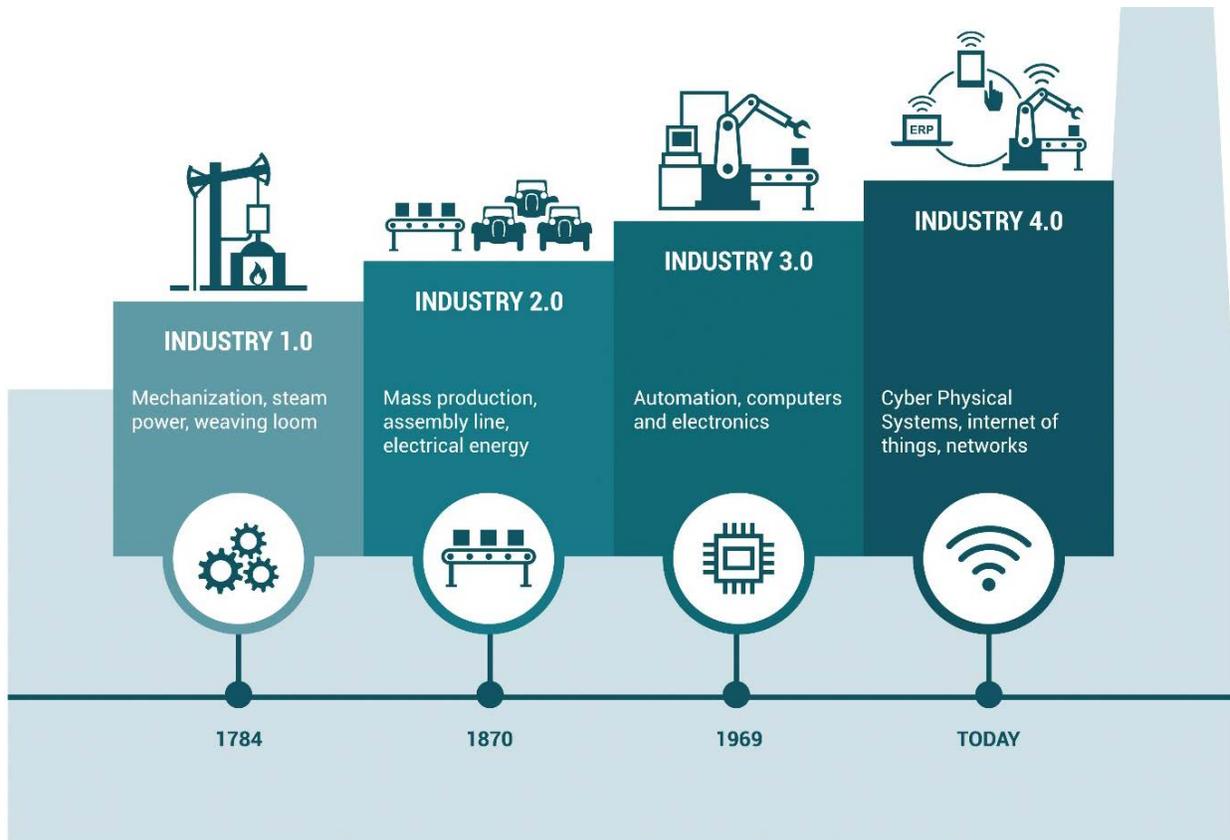


Figure 3. Progression to the Fourth Industrial Revolution and Smart Automation (Source: Adobe Stock Image).

vectors^c is critical to safe, reliable, and sustainable operations of autonomous systems and intelligent ecosystems. Key aspects of assured operations of these systems include public safety, reliability, resilience, cybersecurity, commercial and international partnerships, and continuous objective assessment of state-of-health of intelligent ecosystems. Society can benefit from policies assuring intelligent ecosystem performance, efficiency, and mission effectiveness without adverse effects to lives and property.

These are elements of the current Fourth Industrial Revolution³ shown in context of progression from

^cA threat vector is a means of attacking or degrading system performance or quality of operations.

the First Industrial Revolution (see Figure 3). The power of the Fourth Industrial Revolution technologies is amplified by how they combine and generate innovations. These same technologies also amplify challenges. Widespread adoption of “black box” artificial intelligence (AI) could make autonomous and intelligence ecosystems exceed human capacity to control and understand. Managing these risks requires a new model for governance.

Third-generation complex systems are monitored by comparing metrics and performance to known operational limits or variations from trends. These methods will start to fail in increasingly autonomous systems as the systems can subtly change their behaviors over time or compensate (sometimes

Table 1. Threat Vectors for Intelligent Ecosystems

Threat Vectors	Description
Cyber Attacks	Malicious efforts to subvert a system through software malware or intrusion to command and control a system
Orbital Debris and Collisions	Impacts of satellite debris and micro-meteorites colliding with spacecraft
Space Weather Impacts	Energetic particles from solar flares and coronal mass ejections impinging on space systems affecting electronics
Human Error	Errant commands, programming glitches, design or manufacturing flaws
Sensor Degradation	Change in sensor monitoring characteristics and performance over time affecting measurements and resulting actions
Component Failure	Failures caused by age, excess temperature, excess current or voltage, ionizing radiation, mechanical shock, stress or impact, operating cycle, and many other causes
Radio Interference	Intentional or unintentional impact to system performance resulting from insufficient spectrum management
Unintended Intelligent System Actions	Unintended changes in system performance and actions over time resulting from artificial intelligence and/or machine-learning evolution

erroneously) for internal failures or continuous improvements. Even when an anomaly is detected in one part of the system, that information may not be shared or leveraged by the larger ecosystem. Connected elements might continue to process erroneous data and provide corrupted outputs, causing the issue to further cascade. In these intelligent ecosystems, small abnormalities may spread unchecked resulting in unforeseen downstream impacts. An autonomous system can change its operating environment, which changes

inputs to the system, causing feedback loops that are difficult to track and manage. There are multiple scenarios where time-critical autonomous systems require improved operational assurance.

Consider a hypothetical scenario: During a severe weather event, a space-based instrument in the national environmental intelligence enterprise has an undetected calibration failure. The data generated by the instrument appears nominal by itself but is

not consistent with the other environmental data being collected. Autonomous techniques being developed for warn-on-forecast can incorporate the bad instrument data and fail to provide accurate severe-weather warnings. Improved realtime operational assurance could detect the inconsistencies and discount the bad data or perhaps even detect and reconcile the calibration failure prior to impacting operational performance and mission success.

Another hypothetical scenario: A missile warning system uses machine learning to identify launches to reduce the likelihood of false alarms. A cyber-attack on the system could modify the training data used for the machine learning, resulting in misclassifying actual events as false alarms. This situation could be reconciled using methods to increase assurance of machine learning-based systems by periodically testing their responses to known inputs.

Threat Vectors

There are many threats to consider in maintaining optimal performance and assuring mission success. For space-based systems, threat vectors include orbital debris,⁴ space-weather impacts,⁵ cybersecurity,⁶ human error, sensor degradation, and component failure. An additional type of threat to intelligent ecosystems includes unintended consequences of system evolution associated with artificial intelligence, intelligence agents, and machine learning. Table 1 (above) describes a set of threats to intelligent ecosystems.

Expertise, Technologies, Tools, and Processes

The assurance of effective, reliable, sustained operations of an intelligent ecosystem requires a diverse range of experts and tools. Foremost are experts with knowledge of the design of intelligent agents within the ecosystem, whether they be autonomous vehicles or environmental intelligence systems using space-based Earth observations. Data analysts are needed to manage the massive amounts of data and develop systems to monitor trends and find abnormalities and anomalies. Computer scientists, information systems professionals, and cybersecurity experts are necessary to develop flexible secure frameworks for autonomous systems to operate.

At a minimum, experts and tools need to provide the following functions to assure safe and reliable operations:

- ◆ **Warn** – provide indications to the responsible operators of any anomalous behavior and users of any impact found in the Assess function.
- ◆ **Assess** – determine potential impacts of anomalies in one system to the rest of the ecosystem. In very complex systems, it can often be difficult to determine how one element impacts any final products or safe and reliable operations. An objective is to identify errant

behavior in time to mitigate or avoid cascading effects that ripple through the ecosystem.

- ◆ **Determine Root Cause** – leverage data using analytics and machine learning to determine potential causes of any anomalies.
- ◆ **Mitigate** – minimize the impact of any anomalous behavior through means such as removing erroneous data inputs, activating redundant systems or components, switching to reserve capabilities, shutting down errant commands, or employing fail-safe contingencies.

The tools used for the assurance of intelligent ecosystems should satisfy several requirements. The tools need to verify and validate operations and performance and react in realtime to events. The tools and processes to monitor and manage the performance of autonomous, intelligent systems need to be autonomous as well to keep up with the velocity and volume of the data. A realtime capability implies a continuously streaming data architecture as opposed to conventional batch processing of data. The tools need to identify and address degradations, changes in latencies, and changes across the entire digital ecosystem.^d This implies a broad focus on sensors, networks, and in-line processing, as well as use of models and simulations and end-user validation.

An assessment tool should be reliable with an architecture resilient to failures and redundancy for critical data paths. As data volumes increase, scalability becomes an important requirement. Ideally, system resources should scale with increased throughput. As the components of intelligent ecosystems become more connected,

^dA digital ecosystem is a distributed, adaptive, open socio-technical system, with properties of self-organization, scalability, sustainability, inspired from natural ecosystems.

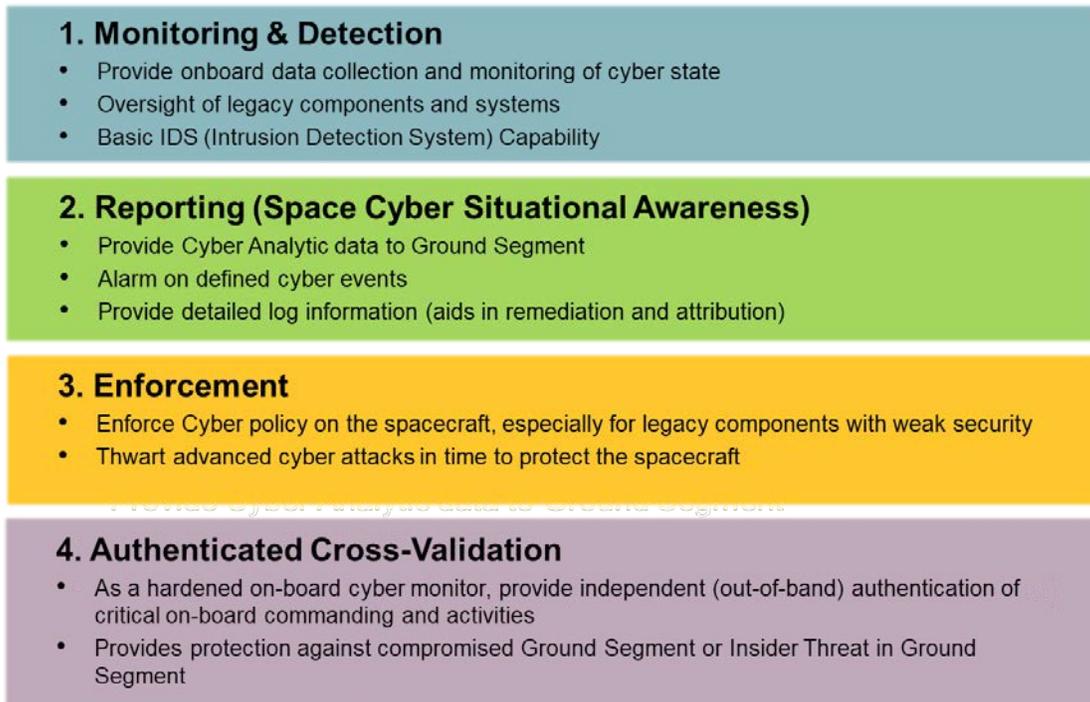


Figure 4. CEASE – Embedded spacecraft cyber defense functions (Source: The Aerospace Corporation).

especially when connected to the IoT, they become increasingly vulnerable to cyberattacks. An assessment capability needs to provide cybersecurity in the form of authorization, Unix-style permissions, and data encryption, in order to operate independently from the system it is assessing. Also, the assessment tool architecture needs to be able to adapt and dynamically embrace and react to change.

Technologies are evolving rapidly and the application of techniques is radically changing many fields. An operations assessment project should be able to start with the existing configuration of a system and add value incrementally, as well as respond quickly to emerging threats. Current best practices in industry emphasize agile development and containerized microservices to implement data systems in a dynamic environment. The Aerospace Corporation is working with the U.S. Department of Transportation to implement a secure Situation Data

Clearinghouse Proof of Concept for collection, fusion, and sharing of realtime operational data for connected and autonomous vehicles within our transportation system. The system is based on containerized microservices operating in a secure cloud environment.

Space asset protection tools and expertise are being deployed to continuously monitor and diagnose state-of-health in complex systems. Initial developments to move operational assurance beyond traditional ‘redline’ or trend monitoring, into leveraging machine learning-based techniques, have shown to be fruitful. The Aerospace Corporation’s Satellite as a Sensor (SAS)⁷ tool uses machine learning to baseline normal operations for a spacecraft. The tool trains on vast amounts of spacecraft telemetry. Once the SAS tool establishes a baseline for normal operations over a range of situations, it can identify anomalies and can distinguish between telemetry data in the normal range and what might be an anomaly, such as a

system malfunction or an attack. It also provides forensic analysis to assist in determination of the root-cause of anomalous behavior. Value has been realized by going beyond monitoring the telemetry of a single spacecraft to gathering inputs from multiple spacecraft. Additional benefits from realtime assimilation of external data sources are being evaluated.

To provide cybersecurity for aerospace systems, Dr. Josh Train, The Aerospace Corporation, notes “We’ve developed a whole series of in-house tools, processes, and techniques to help process your data more effectively. The Cyberspace Operational Environment tool helps us understand all the threats to satellite systems. It helps us determine when data has been compromised or breached. We can apply a whole suite of advanced data analytic tools and techniques to identify trends and anomalies to help protect, repair and recover communications systems.” Figure 4 shows key components of the Cyber Embedded Analytics and Space-based Enforcement (CEASE) cyber defense tool.

The unique aspect of applying artificial intelligence and machine learning to space problems is the stakes are much higher than uses such as commercial social networking and customer intelligence systems. It is considered acceptable in some applications to have AI success rates hover around the 85- to 95-percent mark of what a human can perform. For non-critical applications, this is adequate. However, in the space realm, we frequently need assured operational performance of at least 99.999 percent. This performance threshold requires additional scrutiny of the machine techniques and a deeper level of research than is currently being conducted by many commercial customers. The Aerospace Corporation is building a team of experts with capabilities to further improve performance and reduce risks by understanding each AI technique in depth, and to push the envelope ensuring reliability of techniques for future use in space systems.

Policy Implications

As Klaus Schwab notes in the Fourth Industrial Revolution, “Long-term forecasters warn not to underestimate existential threats if we fail to align values of AI with human values.”³ There is a need for continuous, realtime, sustained knowledge and management of performance and potential threats for assured operations of autonomous systems affecting lives and property. Policies and regulations need to address system configurations with multiple providers, including a spectrum of commercial players, international organizations, and multiple U.S. government agencies. Objective assessments of performance and threats need to be conducted across the systems lifecycle, from concept and architecture to acquisition and development, integration to operations, and even systems disposal. As mentioned above, threats for operations of intelligent ecosystems grow organically over time, based on evolution models ranging from a rigid organizational structure to a loose collection of actors. As systems grow in complexity and autonomy, user communities can become increasingly reliant on the ecosystem for their health and welfare. Consider for example, unmanned aerial vehicles operating in airspace over populated areas, or space habitats accommodating international communities of astronauts. As intelligent ecosystems increasingly provide significant services, governing agencies will be expected to maintain or improve operational cognizance and assurance of safe operations to protect lives and property.

The authors’ research revealed a very limited body of existing policies dealing with managing autonomous systems in the United States or internationally. We also find there is a paucity of regulation governing operations of autonomous systems. Noting that regulation tends to build over time in response to adverse events, there is an understandable hesitation to regulate under the auspices of not stifling innovation and development.

As autonomous systems are utilized within larger ecosystems, decisions will need to be made as to whether internal mechanisms should perform the functions needed to assure safe and reliable operations (warn, assess, determine root-cause, and mitigate), or if external, independent mechanisms are more appropriate to perform those functions. Such accountability needs to be coordinated, regulated, and enforced at a level that governs operation of the entire extended intelligent system. Geographically, intelligent systems can extend from local to regional to global and, for space systems, from low Earth orbit to interplanetary. This will require consideration of governance models associated with a central authority to meet demands and protection of end users. The development will work best if common standards and interfaces are established by, and for, the community. Consortiums of the private sector and government stakeholders are recommended to collaborate on development of standards, policies, and regulation to optimize the efficiency and effectiveness of intelligent ecosystems.

Technical revolution is not new to the aerospace industry. What sets apart the technical revolution in AI is the pace at which it is occurring. The aerospace industry faces a new dynamic at every level in the value chain. Human intervention will not always be feasible for making realtime critical decisions in space and on the ground as AI advances. AI can become a reliable decisionmaker through reinforcement learning and continue to be incorporated into aerospace technologies from innovative small spacecraft to command, control, and communications, to ground stations, and everything in between.

Worldwide and in the U.S., existing statutes governing the use of AI are meager,⁸ and largely promote the use of AI through economic incentives.⁹ Tort law polices the actions of AI developers by penalizing them when something goes amiss with the AI they created.¹⁰ Laws and

policies are made by, and govern the conduct of, humans. A key ingredient for negligence is “foreseeability” and as autonomous systems and AI evolve over time, they can behave in ways unforeseeable by their creators. A consequence of reinforcement learning by AI is humans do not bear responsibility for the decisions made by AI and cannot be found at fault when something goes wrong.

The advantages of AI to aerospace, transportation and other critical industries are manifest, particularly by reducing the human tendency to make biased decisions and by improving safety and performance. It is incumbent on aerospace and other industries to understand the consequences of the use of AI, the technical aspects and the economic, business, ethical, legal, and societal aspects, as well as to ensure there are preventive measures and accountability for the performance of intelligent ecosystems employing artificial intelligence, machine learning, and intelligent agents.

Conclusion

Modern, agile models of governance for verification and validation of system state-of-health are needed to ensure safe operations of autonomous systems affecting lives and property. Sponsors, developers, and operators need to apply expertise, tools, and processes to continuously monitor the performance and threats throughout the lifecycle of autonomous systems and intelligent ecosystems. The intelligent systems lifecycle spans across technologies, designs, acquisitions, program assessments, independent verification and validation, system evolution and assured operations.

Government agencies acquiring and operating intelligent systems benefit from well-characterized components and modules that are tested, low cost, reusable, and easily configurable. The Aerospace Corporation is evolving a structured approach for designing, acquiring, building, testing, integrating,

and assuring operations of modules for autonomous, intelligent systems.

The structured approach includes a “Common Criteria” for assessing vulnerabilities based on a proven framework applied to cybersecurity over the past decade and best practices from other industries including aviation, financial, and social media. Aerospace conducts common criteria assessments coupled with a “train the trainer” approach to allow training initiatives to be conducted in a shorter amount of time to efficiently expand value to serve system acquirers, developers, and operators as needed.

Autonomous machine-learning techniques are based on lifecycles of training, experimentation, and refinement that require significant caretaking and shepherding. Aerospace is working to ensure timely detection and notification of abnormalities such that adversaries and/or faulty processes are not able to cause detrimental flaws in autonomous systems. Informed organizations can ensure well-trained models from one autonomous system are made available to other developers and operators. The Aerospace Corporation is assessing best practices to complement human subject matter experts with machine-learning capacity to deal with high volumes of data, to sustain high values of mission success for autonomous systems over time. There is value in applying a system-of-systems approach to assuring autonomous systems continuously and safely operate to meet their intended mission.

References

- ¹ J. Corrigan, “Three-Star General wants Artificial Intelligence in every New Weapon System” Nextgov, November 2, 2017 <http://www.nextgov.com/defense/2017/11/three-star-general-wants-artificial-intelligence-every-new-weapon-system/142225/>.
- ² International Space Exploration Coordination Group (ISECG), “Global Exploration Roadmap,” January 2018, <https://www.globalspaceexploration.org/>.

- ³ K. Schwab, “Shaping the Fourth Industrial Revolution,” World Economic Forum, 2018. <https://www.weforum.org/focus/shaping-the-fourth-industrial-revolution>.
- ⁴ Aerospace Orbital Debris – <http://www.aerospace.org/publications/policy-papers/commercial-space-activity-and-its-impact-on-u-s-space-debris-regulatory-structure/>.
- ⁵ Space Weather – <https://www.gps.gov/cgsic/meetings/2015/steenburgh.pdf>.
- ⁶ Cyber Attacks – <https://www.forbes.com/sites/greatspeculations/2017/09/21/cyber-attacks-pose-roadblock-for-driverless-cars/#3723da591b6f>.
- ⁷ Dr. C. Tschan and Dr. T. Kashangaki, presentation to the 2016 National Symposium on Sensor and Data Fusion referencing OTR-2016-01139-2 – Intelligent Systems for More Efficient Data Fusion Implementation.
- ⁸ National Science and Technology Council, Executive Office of the President. National Science and Technology Council Committee on Technology. Preparing for the Future of Artificial Intelligence, 17–23. Search results for U.S. legislation governing artificial intelligence, machine learning or autonomous systems over the past 50 years at <https://www.congress.gov/>. S. Hrg. 114-562: The Dawn of Artificial Intelligence. See also House of Commons, Government Office for Science, Artificial Intelligence: Opportunities and Implications for the Future of Decision Making. 17–18. Headquarters for Japan’s Economic Revitalization, “New Robot Strategy.”
- ⁹ Search results for U.S. legislation governing artificial intelligence, machine learning or autonomous systems over the past 50 years at <https://www.congress.gov/>.
- ¹⁰ G. Wisskirchen, et al., Artificial Intelligence and Robotics and Their Impact on the Workplace, 53–61. International Bar Association, 2017, <https://www.ibanet.org/Article/NewDetail.aspx?ArticleUid=012a3473-007f-4519-827c-7da56d7e3509>.

