NASA/TM—2019–220345

Supporting Crew Autonomy in Deep Space Exploration: Preliminary Onboard Capability Requirements and Proposed Research Questions

Technical Report of the Autonomous Crew Operations Technical Interchange Meeting

Shu-Chieh Wu
San Jose State University

Alonso H. Vera
NASA Ames Research Center

July 2019

NASA STI Program…in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM—2019–220345

Supporting Crew Autonomy in Deep Space Exploration: Preliminary Onboard Capability Requirements and Proposed Research Questions

Technical Report of the Autonomous Crew Operations Technical Interchange Meeting

Shu-Chieh Wu
San Jose State University

Alonso H. Vera
NASA Ames Research Center

National Aeronautics and
Space Administration

Ames Research Center
Moffett Field, California

July 2019

Acknowledgements

The Autonomous Crew Operations Technical Interchange Meeting was hosted by the Human Factors and Behavioral Performance Element of NASA's Human Research Program and sponsored in part by the Human Systems Integration Division of NASA Ames Research Center. We thank Tina Holden, Jurine Adolph, Donna Dempsey, Brent Beutter, and Dov Adelstein for inputs on TIM agenda, Angela Carter for event venue planning, Chris Maese and Brian Gore for logistics support, Dominic Hart and Natalia Russi-Vigoya for photography, Phil So for badge and map design, Rick Chen for digital sign design, Patty Lau and Kenji Sotto for registration desk support, Estelle Dodson, Ashcon Nejad, and Brandon Tabaska for Adobe Connect support, Audra Ruthruff, Ben Ruthruff, Emily Ruthruff, and Jennifer Granada for catering support, Jennifer Granada for invitational travel processing, Cody Evans, Alicia Krolak, Swapnil Gaikwad for wallpaper printing, Jon Holbrook and Matt Johnson for facilitating, and last but not least, Diana Acosta, Mario Berges, Dorrit Billman, Dave Blakely, Charlie Dischinger, Dave Foyle, Bonnie John, Parimal Kopardekar, Jessica Marquez, Cynthia Null, Steve Robinson, and Anthony Scroggins for assistance with speaker invitation. We would also like to thank Mary Connors, Tina Panontin, and Lee Stone for comments and suggestions on an earlier draft of the report.

Available from:

NASA STI Program
STI Support Services
Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199

This report is also available in electronic form at http://www.sti.nasa.gov
or http://ntrs.nasa.gov/

Table of Contents

Acronyms and Definitions

3d ....three-dimensional
ACAWS ....Advance Caution and Warning System
ACE ....Assurance, Composed, and Explained
AES................Advanced Exploration Systems
AI ....artificial intelligence
AMO ....Autonomous Mission Operations
AR..................augmented reality
ARIAS ....Assurance Reasoning for Increasinly Autonomous Systems
ASO ....Autonomous Systems and Operations
ASRS ....Aviation Safety Reporting System
ATCS ....Active Thermal Control System
AWS ....antisubmarine warfare
BAS................Building Automation Systems
CAD ....computer-aided design
CAPCOM ......Capsule Communicator
CDRA ....Carbon Dioxide Removal Assembly
CEO ....Chief Executive Officer
CSA................Canadian Space Agency
DC..................damage control
DoD................Department of Defense
DSG ....Deep Space Gateway
DST................Deep Space Transport
ECLSS ....environmental control and life support system
EPS ....electrical power system
ESA................European Space Agency
EVA ....Extravehicular Activity
FCS ....Future Combat System
FDIR ....Fault Detection, Isolation, and Recovery
ft.....................feet
GoG................graph-of-graphs
HEOMD.........Human Exploration and Operations Mission Directorate
hr....................hour(s)
HRP................Human Research Program
HSI.................human-systems integration
IECST ....ISS Exploration Capability Study Team
IFI ....Items for Investigation
IMS ....Inventory Management System
ISHM ....Integrated Systems Health Management
ISS..................International Space Station
IT....................information technology
JAXA ....Japan Aerospace Exploratioon Agency
JPL ....Jet Propulsion Laboratory
LCS ....Littoral Combat Ship
LRR................Logistic Reduction and Repurposing

MCC ....Mission Control Center
MCM.............mine countermeasures
MER...............Mission Evaluation Room
MGV ....Manned Ground Vehicle
min ....minute(s)
NAF ....Navy Availability Factor
NASA ....National Aeronautics and Space Administration
NEEMO ....NASA Extreme Environment Mission Operations
NextSTEP ......Next Space Technologies for Expoloration Partnership
NILM ....non-intrusive load monitoring
OAET.............Office of Aeronautics, Explorations, and Technology
Ops Lab..........Operations Lab
RAPSAP ....Resilience and Portable Sensorimotor Assessment Platform
REALM ....RFID Enable Automatic Logistics Management
RFID ....Radio-Frequency IDentification
RTOP ....Research and Technology Operatin Plan
sec ....second(s)
SME ....subject matter expert
SML ....Suitability for Machine Learning
SPDM ....Special Purpose Dexterous Manipulator
SSC ....Station Support Computer
SUW ....surface warefare
TIM ....Technical Interchange Meeting
TOCA ....Total Organic Carbon Analyzer
TRL................technology readiness level
UV..................unmanned vehicle
VR..................virtual reality
WDA..............Watson Discovery Advisor

Supporting Crew Autonomy in Deep Space Exploration: Preliminary Onboard Capability Requirements and Proposed Research Questions

Technical Report of the Autonomous Crew Operations Technical Interchange Meeting

Shu-Chieh Wu[1] and Alonso H. Vera[2]

Executive Summary

Communication delays are a critical challenge posed by long duration deep space exploration. Space missions historically have relied on an ever-present Mission Control Center (MCC) to direct operations in near real-time. As unanticipated anomalies that defeat fault detection and resolution systems do arise, the lack of real-time communication will significantly weaken what the MCC support represents: a reliable safety net for the flight crew through its deep and diverse areas of expertise and investigative resources. As a consequence, future space vehicles and habitats need to be equipped with capabilities to support the flight crew to operate with little or no ground support. Considerations must be given to vehicle and mission designs that will fortify the traditionally ground-centered safety net and forge new support systems, when communication delays exist.

In August 2018, NASA's Human Research Program, through its Human Factors and Behavioral Performance Element, convened a Technical Interchange Meeting (TIM) on Autonomous Crew Operations at NASA Ames Research Center. The goal of the meeting was to gather input from NASA centers, industry, academia, and branches of the Department of Defense (DoD) to address how intelligent technologies can be applied to augment onboard capabilities to support crew anomaly response. The TIM featured 24 presentations by 29 speakers and hosted a total of 59 attendees, including 43 from 5 NASA centers (Ames, Johnson, Langley, Marshall, and Jet Propulsion Lab) and 4 from the DoD (3 from Army Research Lab and 1 from Naval Postgraduate School), with remaining attendees from academia (e.g., UC Davis, CMU) and industry (e.g., IBM, Siemens).

Discussions were centered around three themes: standards and guidelines, lessons learned in analog environments, and technologies. To help provide a framework for discussion, a concept matrix describing anomaly response processes was created prior to the TIM (Figure 1, page 6). The matrix captures the steps involved (monitoring and detection, diagnosis, solution development and evaluation, solution implementation and verification, resolution documentation) as well as the resources and capabilities required to support these steps (data, knowledge, analysis, synthesis,

[1] San Jose State University; San Jose, California.
[2] NASA Ames Research Center; Moffett Field, California

resource management). A wallpaper size printout of the matrix was utilized at the TIM to solicit attendee inputs along the three themes; the activity garnered 108 submissions of ideas.

Overall, what emerged from TIM discussions was a picture of mismatch between crew anomaly response needs and support that can be provided by existing intelligent technologies. The needs are broad, spanning multiple steps and processes/resources, with many of which lacking support from existing technologies, such as knowledge management throughout the steps of problem solving (especially in resolution documentation) and manpower management. The solutions provided by existing intelligent technologies are specific to the steps/processes that they are designed to support and constrained to solving only problems similar to those that have occurred before. What is lacking from technologies is typically made up by humans, specifically their complex critical thinking, creative problem solving, and domain expertise.

In the end, the TIM highlighted the pressing need to support responses to onboard anomalies during autonomous crew operations, particularly those that have eluded the system tests, inspection, and other assurance processes. Such anomalies can potentially threaten crew and vehicle safety, as well as significantly impact overall operations with additional workload. These fairly rare events are difficult to anticipate and prepare for, given the state-of-the-art in intelligent technologies. This is true even for anomalies that stem from "unknown knowns"—cases in which there is sufficient external information to characterize the problem but the overall pattern fails to be recognized by the problem solver, or in which the internal knowledge needed to solve a problem is held tacitly and potentially accessible by the problem solver but not articulated. It follows that the ability to tackle anomalies lies not only with the availability of relevant information and knowledge but also their accessibility in times of need. To that end, we propose research questions along the following three broad themes:

- How intelligent technologies can help make relevant knowledge and information available?
- How intelligent technologies can help make relevant knowledge and information accessible?
- How intelligent technologies can help support the crew operating as a team in anomaly response processes?

1. Introduction

The success of long duration deep space exploration missions will require fundamental changes in spaceflight operations. Instead of depending on frequent real-time communications with a large Mission Control Center (MCC) ground support team of diverse specialties to provide direction (see Dempsey et al., 2018 for a detailed description of the range of support MCC provides), a flight crew of 4 will need to function more autonomously in response to anticipated reduction in communication quantity and quality as well as unanticipated blackouts. The impact will be felt particularly in dealing with unanticipated, off-nominal situations. A delay or absence of ground support during unanticipated contingencies can become a significant hazard and increase the risk of jeopardizing crew and vehicle safety if there are not sufficient onboard capabilities to assist with troubleshooting and contingency management. The attainment of increased crew autonomy therefore lies critically in the ability to augment the vehicle/habitat with sufficient capabilities so that the 4-person flight crew can, on their own, perform the kind of anomaly response that had previously been done mostly by MCC in the face of the expected communication delays or unexpected blackouts.

As a first step toward identifying onboard capabilities that will enable the flight crew to troubleshoot unanticipated anomalies on their own, including the intelligent technologies that might be a part of those capabilities, NASA's Human Research Program, through its Human Factors and Behavioral Performance Element, convened a Technical Interchange Meeting (TIM) on Autonomous Crew Operations at NASA Ames Research Center on August 6–7, 2018. The goal of the meeting was to gather input from NASA centers, industry, academia, and branches of the Department of Defense (DoD) to address how intelligent technologies can be applied to augment onboard crew capabilities to support vehicle/habitat troubleshooting and maintenance. The focus was on supporting the flight crew in troubleshooting anomalies that are neither outright emergencies (i.e., with well-established response protocols) nor slow-rolling enough that the crew are confident they can wait for the ground to work them. Rather, it was on the types of anomalies that affect critical systems with high uncertainties in consequences if left untended that would require the crew to immediately begin troubleshooting without ground support.

The objective of the present report is twofold. First, it aims to capture and distill ideas that emerged during the TIM, particularly in terms of the support the crew will need for troubleshooting and the technologies that potentially can provide such support. Second, it aims to examine these ideas in light of the processes and needs behind troubleshooting unanticipated anomalies, identify the areas of insufficiencies, and propose topics for future research.

The present report will begin with a background discussion on crew autonomy and introduce a troubleshooting framework to aid discussion. It then continues with a recap of the TIM, focusing on troubleshooting needs and support. It concludes with a set of research questions that need to be addressed in order for the capabilities to be developed and validated.

1.1 Motivation for Crew Autonomy

As human space exploration progresses from low Earth orbit to the Moon and beyond, communication between crews in space and flight controllers on Earth will experience increasing one-way light time delays, from 1.3 sec for the Moon to 22 min for Mars at its maximum distance from Earth (Frank et al., 2013; Love & Reagan, 2013). In response, crews will have to function more independently from mission control on the ground, taking on a more active role in directing, conducting, and planning missions and maintaining systems. Such crew autonomy will require more capabilities be built-in onboard vehicles or habitats, most likely in the form of intelligent systems.

Interactive execution aids, system situational awareness and prognostics, and time-critical diagnostics and decision support—all now provided by mission control personnel and infrastructure on the ground—will be needed in real-time onboard. It follows that teaming of human and machine intelligence will also be essential as many tasks may not be solved by humans or by machines alone.

The idea of having a space vehicle or habitat equipped to handle all tasks currently performed by Mission Control may conjure up the vision of an "MCC in a box." It is true that in enabling crew autonomy, the objective is to load up future space vehicles/habitats with capabilities that are currently provided by the MCC. In reality, however, the end product of a vehicle/habitat enhanced with MCC capabilities will differ from an "MCC in a box" in several major ways. First, the prospective size/complexity of the future vehicle/habitat and the level of technological advancement today will limit the range of possible intelligent support systems in the foreseeable future—in all likelihood it will only be possible to transfer a subset of MCC capabilities onto the vehicle/habitat. Second, the idea of an "MCC in box" often carries with it the assumption that the crew will play a minimal role in addressing anomalies, ignoring the fact that humans are far superior adaptive problem-solvers than any current or foreseeable intelligent technology. Even though many Mission Control capabilities (e.g., system monitoring) may be implemented as onboard automation requiring minimum crew involvement, others will likely appear as intelligent aids supporting crew decision-making and procedure execution. These will require a kind of collaboration similar to that between the crew and the MCC. The questions then become those of identifying what subset of MCC capabilities to provide, determining how best to make them available onboard, and how to promote human-machine teaming.

1.2 Anomaly Response Framework

Anomaly response refers to activities that operators undertake in response to a system fault or a cascading set of system disturbances (Watts-Englert, Woods, and Patterson, see Sgobba, 2018). They commence following the detection and recognition of an anomaly (or a precursor to one) to fulfill broadly one of two functions: troubleshooting (diagnostic search) or response contingency management. According to Watts-Englert and colleagues, the processes of troubleshooting and contingency management do not unfold in a linear sequence but often proceed in parallel and feed into each other. To be able to identify and support autonomous crew's anomaly response needs, it is necessary to understand what processes underlie troubleshooting and contingency management.

Davis and Hamscher (1988) describe troubleshooting as an interaction of prediction and observation, accomplished by solving three subproblems: generating hypotheses by reasoning from a symptom to a set of causes; testing each hypothesis to see which one(s) can account for all available observations; and discriminating those hypotheses that survive testing. Much of the process is intimately tied to understanding the behavior of the system being troubleshot, which is required to advance the process and at the same time is updated and enriched by the outcome. For troubleshooting complex systems such as space vehicles/habitats, it is difficult to develop knowledge (or models) for them. However, as Rasmussen (1985) argues, complexity is not an objective feature inherent in a system but depends on the resolution applied for information search; he notes that, for example, a simple object becomes complex if viewed under a microscope. It follows that, as troubleshooting depends on information, the level at which information is represented and delivered to the troubleshooter can have determining effects.

Contingency management roughly speaking concerns what to do next. Its activities include plan selection, plan modification, contingency evaluation, and safing. Contingency management typically proceeds in parallel to troubleshooting before the nature of the anomaly is known but its options are constantly being reassessed based on troubleshooting outcomes.

To help provide a framework for discussion during the TIM, an anomaly response matrix was created to capture troubleshooting and the underlying processes (Figure 1). The column headings describe the basic steps; the row headings describe the information and knowledge requirements and underlying processes.

	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	+
Data capture, provision, display						
Knowledge acquire, retain, dispense						
Analysis information, situation, outlook						
Synthesis plan, option, procedure						
Resource Management people, time, stuff						
+						

Figure 1. Anomaly response matrix.

2. Highlights from the Technical Interchange Meeting

2.1 ConOps and Role of Intelligent Systems

What role can intelligent systems play in autonomous crew operations? Three speakers shared big picture perspectives borne from their respective backgrounds.

Former NASA astronaut Dr. Stephen Robinson (now Professor of Mechanical and Aerospace Engineering at University of California, Davis) focused on the specific needs of the crew. In his keynote address, Robinson approached the role of intelligent systems in enabling autonomous crew operations by first identifying what the crew needs (or is expected) to be able to accomplish for survival. To that end, he enumerated the attributes of the MCC and the vehicle that are essential to crew survival and mission success. On the MCC side, he regards flight controllers as playing the roles of scientists (with the ability to understand natural laws, limitations, causes and effects), doctors/psychologists (with the ability to assess risks, physiological states, psychological performance), and engineers (with the ability to identify system requirements, evaluate performance data, construct predictive models). On the vehicle side, he recounts the human elements known to contribute to mission success, which are not coincidentally the same ones upon which crew selection is based—intelligence, adaptability, flexibility, creativity, resilience, teaming ability, etc. Robinson thinks intelligent systems have potential in aiding the crew in most of these areas, even those unique to humans. In particular, he sees the role of intelligent systems in helping the crew maintain state awareness (vehicle, environment, team, resource, and themselves) for decision making as well as access operational knowledge for execution.

NASA Senior Scientist for Autonomous Systems Dr. Terry Fong focused on capability. Fong began his talk with an overview of autonomy both as a systems capability as well as a discipline in engineering. Critically, he tried to clarify what autonomy is not—not automation (but often relies on it); not Artificial Intelligence (but may make use of AI methods); and lastly, not necessarily about making machines intelligent, smart or unmanned (but about making systems independent and self-reliant). Even with autonomy, Fong pointed out, systems can include humans as an integral element; for that, human-systems integration and human-autonomy will be an important consideration in system design. As a case in point, Fong went into detail on the plan for the Deep Space Gateway, a cis-lunar orbiting platform that provides habitation, power propulsion (for orbiting), and docking capabilities for Extravehicular Activity (EVA) and logistics cargo, and supports scientific activities. The Gateway is envisioned to have a crew onboard for only 30 days out of a year; the rest of the time (92% of the year), it will operate without crew. Intelligent systems developed for the Gateway will likely be needed to support the crew when they are onboard, help operate the vehicle autonomously when the crew is not onboard, and enable the transitions from crewed to uncrewed and vice versa. Fong is excited about using the Gateway as a test-bed to validate how the current concept of operations could change by, for example, moving toward a hybrid of International Space Station (ISS) and robotic mission operations.

A Capsule Communicator (CAPCOM) during his previous appointment in the Flight Operations Directorate at Johnson Space Center, Mr. Rick Davis from NASA Headquarters focused on a specific operational need—communication. Davis regards intelligent systems as one of fifteen engineering "long poles" that apply to future crewed Mars surface missions. Even though he too envisions that the role of the MCC will evolve from mission control (detailed management of the system) to mission support (advanced strategic mission planning), Davis contends that the ability for astronauts to engage in high-quality and frequent communication with the MCC will remain integral to problem-solving and consequently to mission success. He suggested an out-of-the-box idea to counter communication delays between the Earth and Mars using AI: in his design, each participant engaged in a communication will have a virtual representation constructed using psychological or other profiling techniques. The virtual representation stands in as the local portal and carries on live communications with the distant participant, and another portal system sits in between the two virtual local portals and ties back the time-lagged communications using AI. Davis envisions the technology being tested in steps (text only, voice, and finally video).

2.2 Needs

2.2.1 Insights from the Field

What kind of troubleshooting might be expected on autonomous crew missions? While development of Gateway is still in the early design phase, insights into troubleshooting support needs can be gained from the ISS, the current and best model of extended crew operations in space (Dempsey et al., 2018). With a pressurized volume of 32,333 ft^3 (about that of a Boeing 747), the ISS is the largest human-made body in low-Earth orbit. Its first component was launched in 1998 and the last pressurized module (except for tech demo) was added in 2011. The first resident crew, Expedition 1, arrived in November 2000. It has been continuously crewed and monitored since then.

The ISS has two segments, U.S. and Russian. The U.S. segment includes all the non-Russian partners: the European Space Agency (ESA) module, the Japanese modules operated by Japan Aerospace Exploration Agency (JAXA), and the Canadian robotic systems operated jointly between NASA and the Canadian Space Agency (CSA). The rest is the Russian segment (Dempsey et al., 2018).

The ISS consists of a large, complex web of safety-critical hardware and software components. Its main body is made up of 16 pressurized modules along with 130 other structural and non-pressurized components. Eight miles of wire connects the electrical power system aboard the space station. 100 data networks transfer 400,000 signals (e.g., pressure or temperature measurements, valve positions, etc.) from 350,000 sensors. In the U.S. segment alone, more than 1.5 million lines of software code run on 44 computers to control and monitor the various systems.

The day-to-day operation of the ISS requires 12 ground flight controllers in the main control room (the Front Room) to monitor and attend to crew and station health and safety on a full-time basis. They are supported by many more experts in the "back room" and the Mission Evaluation Room (MER) when off-nominal or anomalous situations arise. These ground teams work in shifts 24/7 to investigate and resolve the issues, involving the crew when necessary to assist in troubleshooting and/or to actively repair or replace hardware. Critically, no engineering data or safety information systems are hosted on the ISS, nor are the data available to the crew. It is primarily the back room flight controllers and MER experts who are responsible for problem investigation/assessment and recovery procedure planning. A recent study of the functions of the decision-making process utilized with current ISS operations found that of the total 14 functions identified, 11 are entirely performed by mission control, with the remaining 3 performed by the flight crew in communication with flight controllers (Tokadlı & Dorneich, 2018). Even though subsystems onboard the ISS (e.g., life support, thermal control) do operate in an automated manner and alert the flight crew to certain faults or off-nominal conditions, the crew have no onboard assistance for problem-solving if communication and interaction with mission control is interrupted or delayed.

Experience with the ISS, as captured in the Items for Investigation (IFI) and other data systems, suggests that it is highly probable that unanticipated off-nominal conditions will occur in space vehicle/habitat systems, especially in the checkout and initial operation phases. This was the focus of the presentation by Dr. Alonso Vera, Chief of Human Systems Division at NASA Ames Research Center. Vera collected and analyzed IFIs prioritized as High Level, which include criticality 1 and 1R components (those that have no redundancy or that reduce redundancy to zero). He found that, since 2000, there had have been about 67 high priority IFIs, with the majority of them (> 65%) occurring early in the operation history of the ISS (2000–2006) and involving every major subsystem. Aside from EVA, the environmental control and life support system (ECLSS) has had the highest rate of occurrence (~20%). Vera drew attention to two IFIs: one involved a fault eventually isolated to the Flow Control Valve in the Active Thermal Control System (ATCS); the other concerned an anomaly on the ECLSS Carbon Dioxide Removal Assembly (CDRA) removal performance. Both required significant amount of time (ranging from two weeks to two months) and data analysis by the MCC to resolve. More importantly, both involved immediate engagement by MCC—troubleshooting the problem quickly to understand how fast it was progressing and what the next possible worse case scenario was. With light time delays, even of just 40 minutes, this first critical step in anomaly diagnosis will likely need to be started by the crew.

The talk by Mr. Steve Hillenius, a UX manager and designer at NASA Ames Research Center, and Mr. Marc Reagan, Mission Director for multiple NEEMO (NASA Extreme Environment Mission

Operations) missions and Station Training Lead in Mission Operations at NASA Johnson Space Center, provided a number of examples of the likely kind of processes involved in troubleshooting off-nominal or anomalous conditions in a space vehicle or habitat. Their examples came from observations of solving unanticipated problems in Aquarius, the undersea research station and habitat analog for the NASA (NEEMO) mission. In their talk, Hillenius highlighted three cases that occurred on the NEEMO 22 mission.

Case 1 concerned troubleshooting a complex payload experiment called Resilience and Portable Sensorimotor Assessment Platform (RAPSAP). The experiment was designed to investigate changes of gait and balance. The experimental apparatus requires many sensors to be added and fine tuned according to the uniqueness of the environment (e.g., CO_2 partial pressure, which is sensitive to the total pressure). As a result, activities often take longer than expected, ranging from 45 minutes at the low end to as long as 3.5 hours in some cases. To mitigate, more time is allotted to accommodate the particular execution conditions in the habitat. Parts are also repackaged differently each time in response to space constraints.

Case 2 concerned troubleshooting a hardware problem in another experiment, mini DNA. The particular issue encountered was that a Surface Pro laptop used in the experiment would not charge and consequently would lose power before sequencing compilation runs were completed. To complicate matters, flow cells used in the experiment was only good for one time use, a finite resource. Additionally, the experiment is sensitive to temperature and time: the sample can go bad if it takes too long to sequence. After consulting with experts on the ground, the root cause was traced to poor connection of the power adapter due to its requiring a very precise placement.

Case 3 concerned replanning a scheduled EVA due to poor weather outlook (choppy seas). A complicating factor was the availability of support divers. Eventually, support on the ground performed real-time replanning, with some planned activities dropped and diver tasks rescheduled.

Each of these troubleshooting cases was not particularly challenging and did not pose safety risks; nevertheless, they serve as good, if simple, analogs for the kind of troubleshooting the crew may need to do autonomously during future deep-space missions. Hillenius noted that much of the mitigation in the three cases required expertise from subject matter experts (SMEs) from multiple disciplines taking into account wide-ranging considerations and decision criteria. Automation could be made to assist with some of these aspects but large open-ended solutions will likely still require the kind of adaptability that only humans can provide. To illustrate, Hillenius and collaborators identified unique task properties in the troubleshooting processes of these three case and screened them through a rubric of 21 task properties for determining Suitability for Machine Learning (SML) proposed by Brynjolfsson and Mitchell (2017). The 21 include properties such as information and interaction needs, output format, level of abstract reasoning involved, and tolerance of ambiguity. Each task property is assigned a scored, on a scale of 1 through 5, based on how easy the task can be performed by machine learning capabilities, with 1 being very difficult, 3 possible, and 5 very easy. Using this rubric, Hillenius and collaborators found that the SML scores for the three cases to fall between 1.8 to 3.5, suggesting that none of them is suitable for machine learning solutions.

Mr. Van Keeping, an aerospace engineer at NASA Johnson Space Center, and Mr. Matt Guibert, Lead of the Human Computer Interaction Group in the Human Systems Integration Division at NASA Ames Research Center, discussed current in-flight anomaly investigation processes with a special focus on the challenges involved in supporting information needs. Van Keeping began with a discussion of the spacesuit water intrusion incident on ISS EVA 23 (July 16, 2013), in which water inadvertently filled a crew member's helmet (1.4 liter), threatening his ability to breathe and requiring early termination of the EVA. Van Keeping only briefly touched on the processes involved in troubleshooting the incident itself, which was essentially a ground-driven process. Instead, he focused on the decision making process that the program management had to embark on to decide whether to temporarily halt all EVAs in order to avoid another incident during the ongoing investigation. The question for the mission became one of identifying safety critical problems that might arise that could not be addressed without EVAs. Keeping walked the audience through the current lengthy, iterative, and largely manual process of collecting information to answer a simple question like "What are all the hazards that rely on EVA activities as a control?".

Following Van Keeping, Guibert discussed recent work in information integration, which bears the objective of getting the right data to the right people at the right time using structured data and persistent links. Structured data can support more complete queries; persistent links are more traversable/navigatable and durable. The end product is akin to embedding meta-knowledge in the data. Even with such a data system, Guibert noted, troubleshooting remains a ground-driven process due to the need to access SMEs and information sources.

As Gateway or any future space vehicle/habitat systems will be constructed of similar existing subsystems, even with newer (or improved) methods and materials, they will likely encounter a similar distribution and rate of occurrence of issues in the early phases of deployment. Existing ISS IFI data shows that the probability of one high priority issue occurring during any 30-day period in the first years of operation is greater than 60%. Put in context of a notional mission to Mars that spans 30 months (12 months round trip between the low-Earth orbit and Mars, 18 months of surface exploration) (Drake, Hoffman, & Beaty, 2010), with the extended distance raising the possibility of communication disruptions and severely limiting the ability of mission control to manage vehicle health, the chance for the flight crew to have to handle high priority issues on their own will be very high.

2.2.2 Insights from Technology Research

Ms. Angie Haddock, Co-Lead of the Advanced Exploration Systems-Autonomous Mission Operations at NASA Marshall Space Flight Center, Ms. Brooke Cannon, a human factors engineer also at Marshall, and Dr. Kerry McGuire, a space human factors engineer at NASA Johnson Space Center, described NASA's efforts to define habitat and vehicle capabilities through technology research that focuses on developing and demonstrating advanced software systems that support various aspects of autonomous crew operations. In particular, they reported work performed under NASA's Autonomous Systems and Operations (ASO) Project funded by Advanced Exploration Systems (AES), related to decision automation methods that utilize artificial intelligence and machine learning technologies to provide capabilities for anomaly detection, fault isolation, failure prognostics, and failure impact determination (Frank & Aaseng, 2016).

One example is the Advanced Caution and Warning System (ACAWS). ACAWS provides automated assistances in two aspects of the Fault Detection, Isolation, and Recovery (FDIR) process: generating the initial diagnosis of the potential sources of a failure, and offering recommendations for appropriate troubleshooting or recovery procedures (Frank et al., 2013). Even though the underlying technologies adopted from Integrated Systems Health Management (ISHM) have advanced to the level where they can provide real-time assistance for FDIR, designing interfaces that can allow human operators to process and understand the ISHM system information rapidly and effectively remains a challenge. McCann, Spirkovska, & Smith (2013) described an effort to integrate ACAWS into NASA's Deep Space Habitat for diagnosing the onboard electrical power system (EPS) with interfaces to support FDIR as well as exploration of "what-if" failure analysis scenarios and training of novice users. Experimental testing found positive responses from both the flight controllers and the crew. The flight controllers noted reduction in workload and need for coordination while the crew noted the improvement in situation awareness and the ability to perform more autonomously from the ground (Frank et al., 2013) .

Another example is the development of an Autonomous Mission Operations (AMO) software that turns over the operation and management of two ISS systems to the onboard crew, the Total Organic Carbon Analyzer (TOCA) and Station Support Computers (SSCs) laptops (Frank et al., 2015). TOCA is a potable water quality analyzer onboard the ISS. Water analysis needs to be done once to twice weekly; each analysis takes approximately three hours and requires several back-and-forth data and voice communications between the Station and the MCC, more if analysis returns anomalous results. Little to no analysis is done by the crew or software onboard. SSC laptops are non-critical crew computers systems used for a wide variety of purposes, from assessing crew's daily mission plan, operational procedures, to crew personal uses (email and entertainment). They are monitored and maintained by flight controllers on the ground. The AMO software changes the turn-taking by analyzing the TOCA data in real-time first before presenting them to the crew, who then use the AMO software to determine if any action is needed. The MCC is contacted only when the crew determine something is wrong and seek recommended actions (Frank, McGuire, Moses, & Stephenson, 2016). The ISS crew expressed appreciation for the increase in situation awareness that the AMO software provides and its ease of use.

One common caveat of both of the ACAWS and AMO systems is that the utility of automated decision aids is limited by the knowledge coded within it (Frank et al., 2016). In other words, they can only assist with troubleshooting problems that can be anticipated in advance. McCann et al. (2013) acknowledge the challenge and danger that "unknown unknowns" pose but suggest that tools like ACAWS can assist crew members in developing contingency procedures. Frank et al. (2016) caution the tendency for users to overtrust imperfect decision aids and urge system developers to ensure that users understand the system's limitations through user interface design and training.

Mr. James Broyan, Advanced Exploration Systems Logistics Reduction Project Manager at NASA Johnson Space Center, described efforts to improve crew autonomy through more automated and streamlined logistics management. The Logistic Reduction and Repurposing (LRR) project, also funded by the AES, aims to identify and develop technologies that reduce logistical mass, volume, and consequently crew time required for logistics management. One of the technologies is the Radio-Frequency IDentification (RFID) Enabled Automatic Logistics Management (REALM) system, which uses RFID technologies for three-dimensional (3D) localization of crew and logistics items (Broyan, Ewert, & Fink, 2014). Currently the ISS uses the Inventory Management System (IMS) to track over 130,000 items stored in approximately 118 cubic meter of usable stowage space; among them only ~3,200 items have RFID tags. REALM aims to automate the process of updating item movements and locations and provide "truth" location of RFID tagged items to save crew time. The plan was to deploy REALM in three phases (Fink et al., 2017). REALM-1, completed in February 2017, repaired the onboard RFID reader/antenna system to provide pervasive and 24/7 coverage. REALM-2 will utilize a robotic free-flyer (Astrobee; Bualat, Barlow, Fong, Provencher, & Smith, 2015) to extend coverage areas and improve location resolution. REALM-3 is planned to integrate readers into drawers/racks to enable smart dense stowage systems.

Technologies supporting different aspects of mission are typically developed in isolation. To mature the technologies to higher technology readiness levels (TRLs), they must be tested in an integrated environment analogous to the one where they will be used. Next Space Technologies for Exploration Partnership (NextSTEP), also part of the AES, is a public-private partnership model that seeks commercial development of deep space exploration capabilities to support extensive human spaceflight missions around and beyond cislunar space. Mr. Bill Othon, Assistant Chief of the Aerosciences and Flight Mechanics Division at NASA Johnson Space Center, is the lead the NextSTEP Ground Test, which aims to support habitat acquisition strategy by developing requirements informed by analysis and test. Gernhardt and colleagues (2018) described the process by which the ground test objectives were derived. It begins with the mapping of the exploration objectives set by the Human Exploration and Operations Mission Directorate (HEOMD) as well as the phase and capability test objectives set by the ISS Exploration Capability Study Team (IECST) onto representative functional requirements for a Deep Space Gateway (DSG). From there, ground test objectives were then defined to evaluate how well different DSG configurations address each of the representative functional requirements.

2.3 Technologies

2.3.1 Automating Systems and Processes with Deep Learning

It is well known that maintenance can occupy a substantial amount of crew time, as has been the case aboard the ISS. The ability to achieve overall mission autonomy depends critically on a ConOp for Gateway where crew manage onboard problems. What technology may help enable crew to autonomously manage the health of the vehicle?

Dr. Mario Berges, Associate Professor of Civil and Environmental Engineering at Carnegie Mellon University, discussed current and next-gen technologies behind smart buildings and the challenges involved in advancing from sensed buildings to autonomous buildings. According to Dr. Berges, Internet-of-Things (IoT) is beginning to enable much more automation in buildings, though not autonomy, because the latter remains difficult to set up. To help explain the difficulty, Berges borrowed the notion of the uncanny valley from the field of robotics. There, the term is typically used to describe an abrupt “dip” in a human observer’s increasing affinity to a humanoid object as it gains likeness to a real human being. Just as the computer generated human begins to feel almost like a real human being, but not quite, the feeling of affinity is suddenly replaced by an uncanny or eerie feeling, causing the level of affinity to dip substantially (Wang, Lilienfeld, & Rochat, 2015). Berges wondered whether a similar kind of uncanny valley exists that separates buildings that have sensed data available from those that can truly utilize sensed data to function autonomously. The difficulty lies with the limitations of current data-driven solutions, specifically machine learning, in extracting useful information from data.

To illustrate, Berges cited two case studies; both concerned inferring the sensed stimuli with respect to what type the sensors were and what they measured. The first one was on Building Automation Systems (BAS), which can help building managers and owners reduce energy consumption. In an ideal framework, a self-managing BAS can be deployed to any building to automatically manage the information processing. That flexibility is enabled by an information mediator layer that handles the integration of heterogeneous information sources and information sharing among three self managing functions—self-recognition (of own components and their configurations so that the needed information can be automatically retrieved), self-monitoring (of the working status of the components), and self configuration (of the information base based on the outputs generated by the other functions). However, because there is little standardization on the format of device metadata (i.e., information that helps contextualizes measurements or control signals sent from/to a device, such as the location within a building, the physical phenomenon being sensed, etc), such a framework must contend with unstructured and inconsistent labels from heterogeneous systems.

The second case study concerned designing non-intrusive load monitoring (NILM) for residential buildings. The objective of NILM is to provide appliance-level energy metering using data from only a whole-house meter (Bergés, Goldman, Matthews, & Soibelman, 2010). There are two general approaches, event-based and event-less. Event-based approaches rely on detecting events (i.e., abrupt changes in power consumption) then classifying them based on appliance signatures, whose definition would require pre-identified labels generated for local features of events. Event-less approaches rely on inferences generated by factorial hidden Markov Models made computationally tractable by first constraining the state space using domain knowledge.

Both case studies, Berges argues, illustrate the importance of domain knowledge, the key in his view to bridging the uncanny valley. Even though data abound in the physical world, it is information derived from this resource that generates value (Bergés, Lange, & Gao, 2018). And the latter process requires significant domain expertise. Berges provided a very clear assessment of what information can be provided by deep learning systems that are taking in building energy and circuit load health. Deep learning systems cannot answer new questions, only the question(s) they were trained on (as neural nets). The interpretation of answers provided by these systems remains reliant on human domain expertise. Furthermore, it remains the case that most building and circuit representations are top-down and therefore poor at supporting bottom-up questions (e.g., what other outlets are on the same circuit as this one?).

Following Berges, Mr. Mark Chung, co-founder and CEO of Verdigris, described the application of smart building technologies in the form of an IoT Energy Meter, a product developed by his company Verdigris that uses machine learning to track building energy consumption for proactive energy management. According to Chung, renewable energy production rate fluctuates, along with it the pricing. Solar energy is cheap when available, expensive when sunlight goes away. Consumption pricing plus peak usage drive up demand charge so that the overall capacity on the grid is maintained. A smart energy meter such as Verdigris' IoT Energy Meter can be trained on how a building is typically operating and build forecast, predictive models of usage. Once trained, it can take a peek into the future and correct peak usage by shutting down part of the system if it senses demand hitting.

Continuing the theme of data representation and interpretation, Mr. Chengtao Wen, a process control engineer at Siemens, described the principles behind Siemens' new approach to quantifying and exposing the assurance and risk of safety-critical systems. The process of system assurance takes place throughout the entire lifecycle of system development (requirement specification, system development, debugging and formal verification, testing, integration) and extends to runtime. The new approach, named ACE (Assurance, Composed, and Explained), takes data (referred to as "artifacts") generated throughout the process (e.g., rule violations encountered during formal verification, traces generated through debugging, forces recorded by physical simulators) and represent them in heterogeneous graph-of-graphs (GoG) to facilitate continuous risk assessment by decision makers (e.g., system engineers, mission commanders). Because critical aspects of system design as well as test results are captured in one place and presented as "risk hotspots" on an assurance "heat map" to users, this approach helps improve traceability to determining root causes.

Following Wen, Dr. Reed Williams, a research scientist at Siemens, described how ACE can be applied to improve product modeling and simulation, using aircraft maintenance as an example. Williams reiterated what turns data into wisdom are contexts and linkages afforded by domain knowledge and understanding. Spacecraft present very well-modeled environments, yet the relevant knowledge is distributed in computer-aided design (CAD) geometry of equipment and local environment, simulation models of equipment and interaction, models of human agents, etc. Williams illustrated how to bring them to bear together in a use case scenario where the goal is to fine tune the thermal load of an engine. Operation of large drives is typically limited due to the need to avoid thermal overload. Because thermal load cannot be measured directly, operation tends to be conservative. Williams showed how to circumvent this common obstacle by running a simulation in parallel to the real system (aka simulation-in-place) and measuring the thermal load using software-based sensors (i.e., soft sensors) (cf. Kadlec, Gabrys, & Strandt, 2009). This approach allows an online-capable simulation model be built based on existing engineering models and constantly calibrated according to sensor information. When combined with an assurance "heat map" introduced by Wen earlier, information from the heat map can be overlaid over an image of the engine using augmented reality to highlight relevant components to be adjusted and tools needed for the adjustment.

2.3.2 Discovery Systems and other Augmenting Technologies

What does it take to augment human capabilities? In four presentations, experts from IBM and NASA Langley provided an in-depth look at the technology, design, and deployment behind cognitive assistant systems based on IBM Watson cognitive computing technology.

Dr. Jeff Kephart, a Distinguished Research Staff Member at IBM Watson Research Center, opened the session by introducing the concept of embodied AI. Rather than a simple Q&A system, embodied AI can have a brain, sensors (eyes, ears), effectors (hands, feet), and even emotional intelligence. It is effectively a software agent that co-inhabits a physical space with people and uses its understanding of what is happening in that space to act as a valuable collaborator on cognitive tasks.

Dr. Bill Murdock, a researcher and computer scientist at IBM Watson Research Center, focused his presentation on how to support a user's information needs. He contended that information needs constitute a positively skewed distribution with a "tall head" and a "long tail." Tall head represents common questions. Because the questions are foreseeable, it is possible to optimize for each information need, provide highly curated responses, and perform with extreme accuracy. Long tail represents rare events/faults. Because they are unforeseeable, it is only possible to optimize for all of such instances together. Consequently, retrieved answers can only be moderately accurate (but often accurate enough), though what is lacking in accuracy may be compensated by providing more answers to a query. Tall head information is amenable to being implemented in conversational systems by listing and enumerating all instances that will lead to a particular piece of information. Long tail information is more suited to be implemented in discovery systems, providing broad coverage of potential answers.

Dr. Jon Holbrook, a cognitive scientist at NASA Langley Research Center, and Dr. Graham Katz, Senior Managing Consultant at IBM, put the discovery systems that Murdock discussed into an operational context. They described the development and demonstration of a Pilot Expert Advisory System based on Watson Discovery Advisor (WDA) technology, an application of the long-tail Discovery type of system. The Pilot Expert Advisory System was billed as a human-autonomy teaming system that monitors and assesses in real-time states of the human, vehicle, and automation systems and links them with external sources of information to provide flight crew with relevant information in anomalous situations. It was designed to be able to answer questions posed by pilots in natural language and find answers in text sources. In building the corpus of expert knowledge that consists both general and domain specific aviation information, unstructured text from Fedeeral Aviation Administration (FAA) publications (regulatory documents and airman's information manuals), relevant incident knowledge from the NASA Aviation Safety Reporting System (ASRS), aircraft-type specific knowledge, as well as NASA select documents were ingested into the WDA system. SMEs were consulted to construct a list of domain-specific terminology for natural language processing and to provide correct answers to domain specific-questions for training machine learning models. Tested against a use-case based on a real incident, the demo system was able to generate hypotheses about possible systems related to a particular fault message and on factors prone to cause that particular fault, with the correct answers listed at the top of candidate hypotheses. However, Katz acknowledged a couple issues that helped put the initial success in perspective. First, technical specifications and formal engineering terminology did not always match up to the colloquial descriptions that flight crew used. Second, it was difficult for the SMEs to think of questions that they do not usually ask; that is, difficult to think beyond "tall head" questions.

Dr. Jeff Kephart wrapped up the session with a presentation that featured several embodied AI prototypes and research projects. He began the presentation with a hypothetical Mars screw scenario in which an embodied AI agent senses an astronaut's behavior (looking worryingly at a gauge) and offers assistance. The exchange is carried out in natural dialogs and requires the agent to be able to sense the immediate physical space (spatial intelligence) and perform a variety of processes according to context (human behavior analysis, emotion analysis, planning, simulation, reasoning, explaining, diagnosis, preference elicitation). Kephart showcased several embodied AI prototypes in the areas of exoplanet exploration, mergers and acquisition, oil and gas field development. The compellingness of the demos notwithstanding, he acknowledged there remain many embodied AI research challenges: sensing and interpreting the user's environment (multimodal adaptive sensor fusion and rich transcription), interacting with the user (spatial AI and contextual interaction and models of self, world, and people), collaboratively executing high-level cognitive functions (e.g., planning, decision-making), building the software/hardware architecture (spanning Edge and Cloud), and measuring and improving the effectiveness of human-agent interactions.

Mr. Victor Luo, General Manager of the Operations Lab at Jet Propulsion Laboatory (JPL), showcased several virtual and augmented reality based tools developed at the Operations Lab (Ops Lab) at JPL that provide assistance in data visualization, procedure execution, and spacecraft design.

In one video, Luo demonstrated how virtual reality (VR) can be used as a platform for effective multi-dimensional immersive data and environment visualization. There, scientists wearing a head-mounted VR headset (Oculus Rift) are placed in a virtual environment constructed based on a 3D dataset acquired by the Curiosity Mars Rover and rendered through a combination of parallax-mapped 2D images and surflet-rendered 3D point clouds (Norris & Davidoff, 2014). A 2013 study conducted by the Ops Lab found the head-mounted display greatly improved mission scientists' understanding of the Curiosity Rover's environment by improving perceptual accuracy (2x for distance estimation, 3x for angular estimation) with minimum training required to use the system.

In another video, Luo demonstrated OnSight, a multi-platform visualization tool that helped scientists and engineers visualize the surface of Mars (Abercrombie et al., 2017). It includes a web-based 2D/3D visualization tool as well as an immersive mixed-reality visualization environment using Microsoft HoloLens. OnSight includes a unique collaboration feature that enables users in different parts of the (physical) world to meet virtually on Mars ("Meet on Mars" sessions) to engage in discussion in a shared spatial context. There, virtual collaboration is enhanced by OnSlight's capability to track where a user is looking and project it as a "gaze ray" from the head of the user's avatar.

Luo also described a NASA project that exploits Microsoft HoloLens for developing software tools to facilitate hardware development (Noor, 2016). ProtoSpace is a 3D spacecraft design tool used by NASA engineers today. The system can superimpose a computer-generated version of a hardware component and project the image as a hologram over existing physical hardware to the user through Microsoft HoloLens. Such kind of augmented reality (AR) visualization allows users to gain better insight into the actual fit of the component in terms of size, shape, and fit. Similar AR technology can also be used to assist space crew with procedure execution. Procedures can be designed and recorded on the ground then be replayed through VR headsets to astronauts on the space station while they perform troubleshooting or repair.

2.4 Implementation

2.4.1 Individual and Team Performance

Dr. Kaleb McDowell, Chief Scientist of the Human Research and Engineering Directorate at U.S. Army Research Laboratory, discussed wide ranging autonomy related research and applications in the U.S. Army. Similar to NASA's vision of future space exploration, the U.S. Army envisions that Future Combat System (FCS) Manned Ground Vehicles (MGVs) to be fast, lightweight, fully operational while moving, and operated with fewer soldiers than current combat vehicles (McDowell, Oie, Tierney, & Flascher, 2007). Due to the smaller vehicle size, it is expected that FCS MGVs will have limited direct vision around the vehicle and rely on indirect vision systems to provide high-fidelity representation of the surrounding area. They are also expected to have complex interfaces, instrument panels, and information rich displays. It is critical that soldier performance is maintained while operating under these conditions.

Clearly, according to McDowell, past manning models based on standardization of roles cannot serve to reduce crew size, for example, from having 16 soldiers operate 4 vehicles (4 per vehicle) to 7 soldiers operate 4 vehicles. Effective operation of FCS will require heterogeneous human-intelligent technology teams flexibly configured with the right mix for each mission. Though highly desirable, the interchangeability of operators could lead to suboptimal team performance constrained by average operator capabilities as well as result in high-performing individuals being assigned positions not utilizing their full potentials if implemented without accounting for individual differences. In order to enhance emergent team properties contributing to effective team performance, DeCostanza and colleagues (2018) argue that it is necessary to provide technologies to enhance individual team member performance and facilitate the interactions and interdependencies between heterogeneous members of human-agent teams. DeCostanza et al. lay out three broad and intertwined areas of associated technical and scientific challenges: 1) technologies that can adapt to individual differences for the purpose of enhancing overall team performance; 2) technologies that can adapt to the dynamics of tasks and environmental contexts and team members; and 3) training for mutual adaptation and complex teaming.

Dr. Patrick Lincoln, Director of the Computer Science Laboratory at SRI International, approached autonomous crew operations from the perspective of assuring the safety of increasingly autonomous systems where the size of the crew is reduced. He contends that automated systems must operate as part of a human-agent-robot team; that is, a member of the crew. Drawing from the final technical report for NASA project "Assurance Reasoning for Increasingly Autonomous Systems (ARIAS)" (Alves et al., 2018) performed in the context of reduced crew operations in aviation, Lincoln noted four "must-have" qualities of such systems:

- Never Give Up: Automated systems must continue to function sufficiently even after design assumptions are violated. This quality is akin to graceful degradation or extensibility, where a system brings extra adaptive capacity to bear in the face of surprising events challenging its boundaries (Woods, 2015).
- Pervasive Monitoring: There need to be runtime monitoring of subsystems and systems within a system-of-systems context according to explicit top-level safety and mission goals. There also need to be a language/logic for expressing the rules under the monitors and means to communicate to other systems and humans when monitors detect issues.
- Explainable Implementation: There need to be ways to enable humans to understand and gain confidence in automated systems/teammates.
- Rare Conditions: There need to be ways to handle rare conditions that offer no to little data to have trained the automated systems (or humans).

Lincoln then illustrated some of these qualities in technologies of potential utility for autonomous crew operations. The technology behind Siri was born out of SRI's pioneer research in AI, human-computer interaction, and software agents. Today's next generation virtual personal assistants have gone beyond conversation and combine visual and other sensory information to understand and respond to emotion. Combined with augmented reality, virtual personal assistants can turn into mentors and provide interactive step-by-step procedures in diagnosis and repair. However, as devices become more capable and intelligent, they also tend to become more complex and unintelligible. Lincoln argues that trust or trustworthiness is the key enabler to achieving collaborative autonomy in human-agent-robot teams. Many of the barriers to appropriate trust have to do with the fact that autonomous systems are inherently complex and evolve overtime, and the dynamics of the environment makes autonomy unpredictable. To bridge and fill the explainability gap, it is important to characterize the state of the autonomous system in ways that align with human interpretation, to provide meanings of numerical data and decision rationales, and to characterize uncertainty with confidence levels.

2.4.2 Manpower

Discussions of technologies often focus on what capabilities they provide and rarely on what is required to harness the capabilities, yet it is the latter that determines the ultimate success (or failure). Case in point, autonomous crew operations will undoubtedly require a slew of technologies to enable capabilities new both to the vehicle/habitat and the flight crew, particularly for troubleshooting during emergencies. How to determine whether the crew of four will be able to use them effectively at times of need? The issue of manpower is a novel one to space operations that have traditionally relied on (and benefited from) access to near limitless real-time ground support but a central and crucial one to the Navy. In her presentation, Dr. Nita Shattuck from Naval Postgraduate School helped lend support to the issue of manpower by describing a case study based on the Littoral Combat Ship (LCS).

The LCS is a relatively small and agile Navy surface ship specifically designed to operate in the littoral (near shore) area not accessible to Navy cruisers and destroyers. The LCS is a focused-mission ship, equipped to perform one primary mission at any given time; primary missions include antisubmarine warfare (ASW), mine countermeasures (MCM) and surface warfare (SUW) against small boats (including so-called “swarm boats”). The LCS achieves its versatility through modular “plug and fight” mission packages, including unmanned vehicles (UVs); the ship’s mission orientation is changed by swapping out its mission package (O’Rourke, 2014).

The LCS is developed by two industry teams and therefore comes in two different designs. The Freedom class design, developed by Lockheed, is based on a steel semi-planing monohull with an aluminum superstructure, while the Independence class design, developed by General Dynamics/Austal, is based on an all-aluminum trimaran hull. The two designs also use different built-in combat systems (i.e., different collections of built-in sensors, computers, software, and tactical displays).

In 2001, the Navy began an effort referred to as the optimal manning initiative to reduce crew sizes aboard various legacy surface and amphibious ships (United States Government Accountability Office, 2017). The LCS employs automation to achieve a reduced-sized crew. The aim was to achieve a core crew size of 40 sailors. With the additional sailors as needed to operate the ship’s aircraft and mission packages, a total crew of about 88 sailors would be needed, compared to more than 200 for the Navy’s legacy frigates and about 300 (or more) for the Navy’s current cruisers and destroyers.

Unfortunately, both LCS developments have been plagued with design and operational issues. During sea trials, Freedom-class ships suffered repeated engine failures and Independence-class hulls exhibited massive corrosion and transmission failures, necessitating design modifications for both classes. Several crew errors during operations have resulted in significant repairs. These problems caused the Navy to conduct an engineering stand down of all LCSs in September 2016 to assess and mitigate systemic deficits (LaGrone, 2016). A Government Accountability Office investigation was also conducted (United States Government Accountability Office, 2014). Both found that crew training was insufficient and the Navy ordered that every sailor be retrained. It was also found that the core crew of 40 sailors and officers were too few to safely operate the ship without overworking personnel. One consequence of the over-reliance on technology, as Shattuck noted in an anecdote, was that IT (information technology) support technicians became the real-world MacGyver and “Go-To” problem solver for multiple ship casualties (communication network issues, waste system, flooding from hull damage, engine fire). Eventually, the complement was increased to 70 in 2016 (United States Government Accountability Office, 2017). Moreover, because ship operation proved so demanding, six LCS—three of each type—are now dedicated solely to training new crews and another four to testing.

In light of the troubled operation history of the LCS, the objective of Shattuck’s case study was to investigate what the right number and correct composition of crew is for the workload required (Shattuck & Matsangas, 2015; Shattuck, Matsangas, Seagren, & Meredith, 2016). Conventional manpower analysis captures routine duties and events; level of manning is typically determined using the average. Critical phenomena are infrequent but carry dire consequences. How does a system manned according to the average respond to transient phenomena? To answer that question, Shattuck developed three workload models of the LCS crew based on the IMPRINT Pro Forces Module. The basic underlying concept is that crewmembers spend all of their time in some sort of “planned” activities/events, i.e., the ones that typically occur in in the ship’s daily schedule. The planned activities are periodically interrupted by unforeseen events and emergencies (i.e., unplanned events). The three models had increasing levels of operational realism and complexity. The first, baseline model consisted planned activities and some regularly occurring unplanned events. The second model incorporated some irregularly occurring unplanned events. The third model further incorporated "black swans," damage control (DC) events that involved all crew, 12–24 hours in duration. Shattuck found that even under the baseline model, watchstanders worked on average 2.6 hr/day more than the Navy Availability Factors (NAF) daily duty hour provision. Under the second model, engine, gas turbine system techs, and electrician's mates had the highest average daily workload. Under the third model, Shattuck found significant sleep loss and excessive sustained wakefulness; about 30 crew members did not sleep for over 40 hours. Moreover, crew responded mainly to the major events and only critical watches could be maintained, indicating a lack of adequate spare capacity.

Even though many problems of the LCS can be attributed to human-systems integration (HSI) related issues—modernized interface found unusable by the operators, limited design review by HSI professionals, systems overdesigned for its purpose, incomplete training, and consequential operator fatigue and exhaustion over operation, there are manpower specific issues as well. For them, Shattuck highlighted two recommendations from U.S. Navy's Strategic Readiness Review released in December 2017. One is to establish a process to measure the true workload of ships' crews, both periodically and after upgrades and modernizations, to determine if manpower models adequately predict personnel requirements at sea and in port. The other is to adjust ship manning levels to allow for adequate crew rest, performance of extraneous and collateral duties, and training that occurs while onboard ship, and to include some excess capacity.

2.5 Nuggets of Gold (and Lumps of Coal): Ideas from Wallpaper Submissions

Over 100 submissions were received with ideas that roughly fall under one of three themes: information needs, technology, and human systems integration. These are captured in Appendix A and summarized below:

- Ideas on information needs concern acquisition (sensors, knowledge capture), analysis (monitor/detection analysis), and provision (brain books, spacecraft Siri).
- Ideas on technology include input techniques (eye tracking, facial recognition), decision aids (assumption tracking, case-based reasoning, risk assessment of decision model/simulation, what-if scenario generation), and implementation (reconfigurable systems, multi-level automation).
- Ideas on human-systems integration capture common concerns over autonomy (explainability, mental model, trustworthiness, situational cognition, resilience), authority (rejection of solution by humans), and use (perceive and monitor human actions).

3. Path to Crew Autonomy

The Human Research Program (HRP) focuses on applied research necessary to understand and reduce spaceflight human health and performance risks. For autonomous crew operations, the focus is on risks related to onboard crew capabilities necessary to enable autonomous crew missions, as outlined by Human Factors and Behavioral Performance Element Scientist Dr. Tom Williams. Specifically, the risk of:

- an incompatible vehicle/habitat design
- inadequate mission, process, and task design
- inadequate design of human and automation/robotic integration
- inadequate human-computer interaction
- performance errors due to training deficiencies

Dr. Kara Latorella described a formal and systematic effort to review existing NASA agency-level standards and design guidelines for compiling Gateway-specific human systems integration plan and requirements. Dr. Kritina Holden described a complementary effort within the HRP to survey standards and guidelines from other government agencies and industry to augment NASA standards and guidelines. The outputs of these two efforts will address the implementation of capabilities. The present report focuses on the capabilities per se, as well as how the capabilities are to be realized in part through the reduction of those specific risks.

3.1 Preliminary Onboard Capability Requirements

As a rough measure, Figures 2 and 3 respectively map the troubleshooting needs and available support technologies that emerged from the TIM presentations. One immediate impression from comparing the two figures is that the needs tend to be broad and spanning multiple steps and processes while the technologies tend to provide point solutions. The maps also show several areas of need under-covered by existing technologies, such as the need for knowledge management throughout the steps of problem solving particularly resolution documentation, and the need for manpower management.

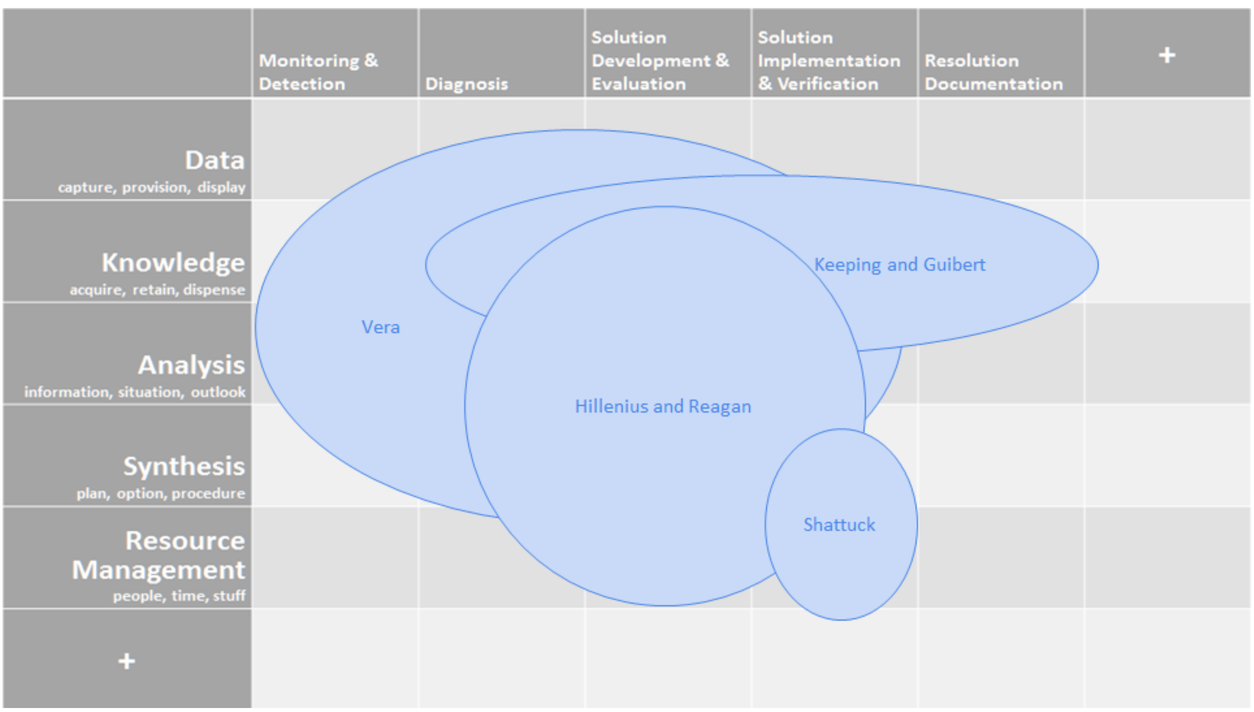

Figure 2. Troubleshooting needs highlighted in TIM presentations.

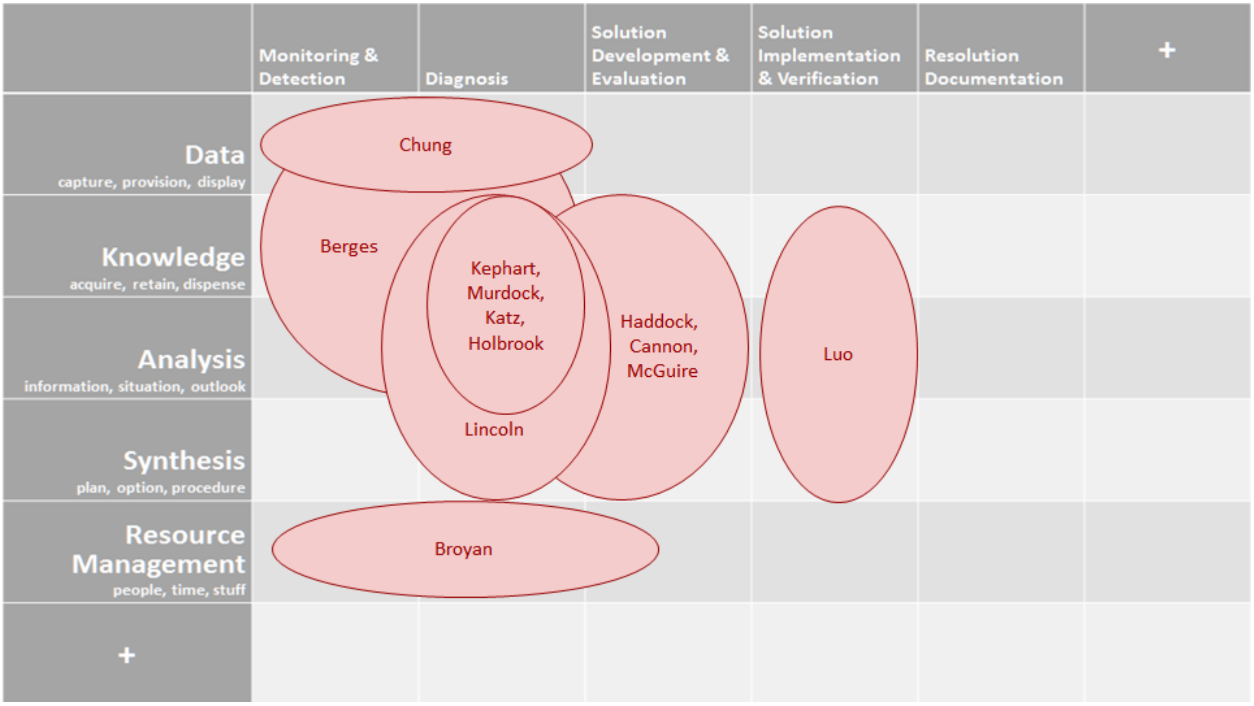

Figure 3. Technologies highlighted in TIM presentations.

3.2 Proposed Research Questions

In the end, the TIM highlights the pressing need for research that will inform the development of human-systems capabilities for supporting the flight crew in executing complex autonomous missions including responding to unanticipated anomalies. Specifically, it calls for research on how best to enable intelligent technologies to complement and enhance human capabilities. The desire to use intelligent systems to assist human operators in aerospace operations, particularly fault management, has been around for several decades. In 1990, as part of a Research and Technology Operating Plan (RTOP) for the Artificial Intelligence Division of the Office of Aeronautics, Explorations, and Technology (OAET), Malin, Schreckenghost, Woods, Forbus and colleagues initiated a multi-year interdisciplinary study to provide guidance in designing intelligent systems to be effective team players in flight operation support. The result of that effort were detailed characterizations of fault management operations and information needs based on observations, interviews, and flight control documents, as well as identified design guidance and outstanding research issues for supporting coordination and management of intelligent systems, fault management processes, and information management (Malin et al., 1991). Many—if not most—of the research issues identified by Malin and colleagues remain unsolved, and their importance is amplified in the severe limitation of manpower among autonomous crews as well as the need to coordinate time-critical activities amid communication delays.

As difficult as it is to prepare to handle unanticipated, potentially high-consequence anomalies, many such occurrences can be resolved with enough time and understanding. The reason is that most anomalies are in fact “unknown knowns,” cases where there is sufficient external information to foresee the problem but the information may be distributed in space (across different data systems and formats) and in time (requiring analyses to coalesce and synthesize). Or, the overall pattern fails to emerge prior to the occurrence for reasons of poor communication, hidden assumptions, etc. It is also possible that the knowledge (general and domain-specific) needed to solve a problem is held tacitly and potentially accessible by the problem solver but fails to be retrieved due to various reasons (Sutcliffe & Sawyer, 2013).

With their unique capabilities in performing complex critical thinking and constructing integrated domain knowledge, humans have proven to be the ultimate problem solver. We envision that intelligent technologies can be utilized to complement and enhance human capabilities by making existing, relevant information and knowledge known—available and accessible—to the crew in a timely fashion and in forms and amounts manageable with their limited manpower. In the following sections we propose research questions whose answers can help realize that vision.

3.2.1 How Intelligent Technologies Can Help Make Relevant Knowledge and Information Available

Crew troubleshooting and contingency management requires information; in autonomous operations, such information must be reside on-board with the crew instead of supplied from the ground. Here we use system information to include both real-time telemetry data as well as device and system knowledge. Much of troubleshooting involves observing and predicting how a system behaves under different conditions (Davis & Hamscher, 1988). For complex systems, this process benefits greatly from an understanding of the interdependencies between different parts of a system. Some tools that support the capture and documentation of system variables and associated behaviors already exist, such as design structure matrices (Browning, 2001; Steward, 1981). However, they are designed to support system development and not anomaly diagnosis and response. The latter conceivably requires not only an understanding of how a certain component should behave given a particular input or output but also the implication of that behavior within and beyond the system in question. Research is needed to identify what aspects of system information (from both the vehicle and habitat) are most critical and thus must be captured and made available to the crew for the purpose of anomaly diagnosis and response.

Even with the right kind of system information presented the right way, domain-specific knowledge and expertise will remain critical for interpreting the information. With a planned crew size of four, determining the right composition of the crew will be critical to assuring that all needed areas of knowledge and expertise will be available onboard. Part of the answer will depend on what types of knowledge and expertise can only be (or are best) acquired beforehand through formal academic education and training, and what types of knowledge and expertise can be acquired through on-the-job or just-in-time training. It could be the case that procedural knowledge can be embedded in aiding technologies and dispensed through augmented reality visualization (e.g., Noor, 2016). Research is needed to systematically survey the areas of domain knowledge and expertise future flight crew should have available to them and the different possible ways to use intelligent technologies to make them available.

3.2.2 How Intelligent Technologies Can Help Make Relevant Knowledge and Information Accessible

Research is also needed to determine how best to represent and display system information to the user with respect to a range of specific contexts (e.g., diagnosis, procedure execution, etc). For example, Berges illustrated how end users tend to seek information from the bottom-up (e.g., What other outlets are on the same circuit as this one?) while system information is typically represented and presented top-down according to the way the system is decomposed and developed. It should be noted that the most suitable way to present system information may vary according to the user and situation; there may not be a one size that fits all. For example, users of different levels of domain-specific knowledge and expertise may be best served with information presented at different levels of abstraction. The level of time criticality when the information is being sought should also be considered; it may be the case that in some situations less detailed information is more (suitable).

3.2.3 How Intelligent Technologies Can Help Support the Crew Operating as a Team in Anomaly Response Processes

The space habitat and vehicle currently being planned, Gateway and Deep Space Transport (DST), present a unique challenge to the issue of manpower. For most earth- or air-bound systems, manpower needs are derived from analyzing a mission's operational requirements, capabilities, and environments, as well as operator workload, productivity constraints, and other factors (cf. United States Government Accountability Office, 2017). Spaceflight has additional constraints that drive to very small crew sizes; for example, Gateway and DST will have a crew size (i.e., manpower) of four. Though the exact composition of crew expertise remains to be specified, the crew will certainly consist of astronauts of diverse expertise. Research is needed to determine how best to pool together human expertise distributed in team members and information distributed across systems efficiently and effectively to support anomaly response.

In addition, it is an empirical research question whether the conventional methods of deriving manpower estimates by going from operation needs to operator competencies (expertise, productivity, and workload) can be utilized in reverse to derive operation capabilities (hardware, software, training) from manpower for spaceflight. Research is also needed to address "black swan" situations (i.e., catastrophic surprises). Shattuck suggested that one solution for black swan type situations is to include some excess manpower. In a condition where manpower is fixed, how might excess be achieved operationally?

4. Concluding Thoughts

Increasingly on the ISS the crews have been provided with tools that enable them to function more autonomously, such as crew self-scheduling tool Playbook (Marquez et al., 2017) and the Special Purpose Dexterous Manipulator (SPDM) "Dextre" (Coleshill et al., 2009). However, it may be argued that the crews are never truly autonomous or MCC-independent until there are sufficient vehicle capabilities to support the crew to investigate and resolve anomalies on their own. With this TIM we took the first step in identifying what capabilities are needed and what intelligent technologies might help provide those capabilities, as well as in identifying fundamental research questions that need to be addressed before the technologies can be successfully applied. We recognize that there remain many challenges in realizing the vision of autonomous crew operations using intelligent technologies, chief among them how to effectively introduce human-systems integration requirements into the system develop process (Pew, 2008). It is hoped that this TIM and the discussion stimulated will assist in the design of future NASA vehicles and habitats and prevent having to relearn lessons of the LCS.

References

Abercrombie, S. P., Menzies, A., Luo, V., Winter, A., Clausen, M., Duran, B., … Lidawer, A. (2017, December). *OnSight: Multi-platform Visualization of the Surface of Mars*. Poster presented at the 2017 American Geophysical Union Fall Meeting, New Orleans, LA. Retrieved from https://agu.confex.com/agu/fm17/mediafile/Handout/Paper246353/ED11C-0134-onsight-agu-web.pdf

Alves, E., Bhatt, D., Hall, B., Driscoll, K., Murugesan, A., & Rushby, J. (2018). *Considerations in Assuring Safety of Increasingly Autonomous Systems: Final Technical Report for NASA Project Assurance Reasoning for Increasingly Autonomous Systems (ARIAS)*. Retrieved from https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20180006312.pdf

Bergés, M., Goldman, E., Matthews, H. S., & Soibelman, L. (2010). Enhancing Electricity Audits in Residential Buildings with Nonintrusive Load Monitoring. *Journal of Industrial Ecology*, *14*(5), 844–858. https://doi.org/10.1111/j.1530-9290.2010.00280.x

Bergés, M., Lange, H., & Gao, J. (2018). Data-Driven Operation of Building Systems: Present Challenges and Future Prospects. In I. F. C. Smith & B. Domer (Eds.), *Advanced Computing Strategies for Engineering. EG-ICE 2018. Lecture Notes in Computer Science, vol 10864* (pp. 23–52). Cham, Switzerland: Springer International Publishing.

Browning, T. R. (2001). Applying the design structure matrix to system decomposition and integration problems: a review and new directions. *IEEE Transactions on Engineering Management*, *48*(3), 292–306. https://doi.org/10.1109/17.946528

Broyan, J. L., Ewert, M. K., & Fink, P. W. (2014). Logistics Reduction Technologies for Exploration Missions. *AIAA SPACE 2014 Conference and Exposition*. Presented at the AIAA SPACE 2014 Conference and Exposition, San Diego, CA. https://doi.org/10.2514/6.2014-4334

Brynjolfsson, E., & Mitchell, T. (2017). What can machine learning do? Workforce implications. *Science*, *358*(6370), 1530–1534. https://doi.org/10.1126/science.aap8062

Bualat, M., Barlow, J., Fong, T., Provencher, C., & Smith, T. (2015). Astrobee: Developing a Free-flying Robot for the International Space Station. *AIAA SPACE 2015 Conference and Exposition*. Presented at the AIAA SPACE 2015 Conference and Exposition, Pasadena, California. https://doi.org/10.2514/6.2015-4643

Coleshill, E., Oshinowo, L., Rembala, R., Bina, B., Rey, D., & Sindelar, S. (2009). Dextre: Improving maintenance operations on the International Space Station. *Acta Astronautica*, *64*(9), 869–874. https://doi.org/10.1016/j.actaastro.2008.11.011

Davis, R., & Hamscher, W. (1988). Chapter 8 - Model-based Reasoning: Troubleshooting. In H. E. Shrobe & the American Association for Artificial Intelligence (Eds.), *Exploring Artificial Intelligence* (pp. 297–346). https://doi.org/10.1016/B978-0-934613-67-5.50012-5

DeCostanza, A. H., Marathe, A. R., Bohannon, A., Evans, A. W., Palazzolo, E. T., Metcalfe, J. S., & McDowell, K. (2018). *Enhancing Human–Agent Teaming with Individualized, Adaptive Technologies: A Discussion of Critical Scientific Questions* (Technical Report No. ARL-TR-8359; p. 38). Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/1051552.pdf

Dempsey, R. C., Contella, D. E., Korth, D. H., Lammers, M. L., McMillan, C. R., Nelson, E., … Van Cise, E. A. (2018). *The International Space Station: Operating an Outpost in the New Frontier*. Retrieved from https://www.nasa.gov/feature/new-nasa-e-book-offers-inside-look-at-space-station-flight-controllers

Drake, B. G., Hoffman, S. J., & Beaty, D. W. (2010). Human exploration of Mars, Design Reference Architecture 5.0. *2010 IEEE Aerospace Conference*, 1–24. https://doi.org/10.1109/AERO.2010.5446736

Fink, P. W., Kennedy, T. F., Rodriguez, L., Broyan, J. L., Ngo, P. H., Chu, A., … Gummeson, J. (2017). Autonomous Logistics Management Systems for Exploration Missions. *AIAA SPACE and Astronautics Forum and Exposition*. Presented at the AIAA SPACE and Astronautics Forum and Exposition, Orlando, FL. https://doi.org/10.2514/6.2017-5256

Frank, J. D., & Aaseng, G. B. (2016). Transitioning Autonomous Systems Technology Research to a Flight Software Environment. *AIAA SPACE 2016*. Presented at the AIAA SPACE 2016, Long Beach, CA. https://doi.org/10.2514/6.2016-5530

Frank, J. D., Iverson, D., Knight, C., Narasimhan, S., Swanson, K., Scott, M. S., … Moses, H. (2015). Demonstrating Autonomous Mission Operations Onboard the International Space Station. *AIAA SPACE 2015 Conference and Exposition, AIAA SPACE Forum*. Presented at the AIAA SPACE 2015 Conference and Exposition, AIAA SPACE Forum. https://doi.org/10.2514/6.2015-4646

Frank, J. D., McGuire, K., Moses, H. R., & Stephenson, J. (2016). Developing Decision Aids to Enable Human Spaceflight Autonomy. *AI Magazine*, *37*(4), 46–54. https://doi.org/10.1609/aimag.v37i4.2683

Frank, J. D., Spirkovska, L., McCann, R., Wang, L., Pohlkamp, K., & Morin, L. (2013). Autonomous mission operations. *2013 IEEE Aerospace Conference*, 1–20. https://doi.org/10.1109/AERO.2013.6496927

Gernhardt, M. L., Beaton, K. H., Chappell, S. P., Bekdash, O. S., & Abercromby, A. F. J. (2018). Development of a ground test & analysis protocol for NASA's NextSTEP phase 2 habitation concepts. *2018 IEEE Aerospace Conference*, 1–27. https://doi.org/10.1109/AERO.2018.8396513

Kadlec, P., Gabrys, B., & Strandt, S. (2009). Data-driven Soft Sensors in the process industry. *Computers & Chemical Engineering*, *33*(4), 795–814. https://doi.org/10.1016/j.compchemeng.2008.12.012

LaGrone, S. (2016, September 5). Navy Ordered Littoral Combat Ship Engineering Stand Down. Retrieved October 9, 2018, from USNI News website: https://news.usni.org/2016/09/05/navy-ordered-littoral-combat-ship-engineering-stand-retraining-ongoing

Love, S. G., & Reagan, M. L. (2013). Delayed voice communication. *Acta Astronautica*, *91*, 89–95. https://doi.org/10.1016/j.actaastro.2013.05.003

Malin, J. T., Schreckenghost, D. L., Woods, D. D., Potter, S. S., Johannesen, L., Holloway, M., & Forbus, K. D. (1991). *Making intelligent systems team players: Case studies and design issues. Volume 1: Human-computer interaction design* (NASA Technical Memorandum No. 104738). Retrieved from https://ntrs.nasa.gov/search.jsp?R=19910021399

Marquez, J. J., Hillenius, S., Kanefsky, B., Zheng, J., Deliz, I., & Reagan, M. (2017). Increasing crew autonomy for long duration exploration missions: Self-scheduling. *2017 IEEE Aerospace Conference*, 1–10. https://doi.org/10.1109/AERO.2017.7943838

McCann, R. S., Spirkovska, L., & Smith, I. (2013). Putting Integrated System Health Management to Work: Development of an Advanced Caution and Warning System for Next-Generation Crewed Spacecraft Missions. *AIAA Infotech@Aerospace (I@A) Conference*. Presented at the

AIAA Infotech@Aerospace (I@A) Conference, Guidance, Navigation, and Control and Co-located Conferences (AIAA 2013-4660), Boston, MA. https://doi.org/10.2514/6.2013-4660

McDowell, D. K., Oie, D. K. S., Tierney, T. M., & Flascher, D. O. M. (2007). Addressing Human Factors Issues for Future Manned Ground Vehicles (MGVs). *Army AL&T*, *January-March*, 20–23.

Noor, A. K. (2016, October 1). The Hololens Revolution. *Mechanical Engineering Magazine*, *138*(10), 30–35.

Norris, J., & Davidoff, S. (2014). NASA Telexploration Project demo. *2014 IEEE Virtual Reality (VR)*, 183–184. https://doi.org/10.1109/VR.2014.6802112

O'Rourke, R. (2014). *Navy Littoral Combat Ship (LCS) Program: Background and Issues for Congress* (p. 104) [Congressional Research Service Report for Congress]. Congressional Research Service.

Pew, R. W. (2008). Some New Perspectives for Introducing Human-Systems Integration into the System Development Process. *Journal of Cognitive Engineering and Decision Making*, *2*(3), 165–180. https://doi.org/10.1518/155534308X377063

Rasmussen, J. (1985). The role of hierarchical knowledge representation in decisionmaking and system management. *IEEE Transactions on Systems, Man, and Cybernetics*, *SMC-15*(2), 234–243. https://doi.org/10.1109/TSMC.1985.6313353

Sgobba, T. (2018). Chapter 13 - Human factors in mission control centers. In B. Kanki, J.-F. Clervoy, & G. M. Sandal (Eds.), *Space Safety and Human Performance* (pp. 569–620). https://doi.org/10.1016/B978-0-08-101869-9.00013-3

Shattuck, N. L., & Matsangas, P. (2015). *Optimization of shipboard manning levels using IMPRINT Pro Forces Module* (Technical Report No. NPS-OR-15-008). Retrieved from Monterey, California. Naval Postgraduate School website: https://calhoun.nps.edu/handle/10945/47006

Shattuck, N. L., Matsangas, P., Seagren, C., & Meredith, I. (2016). *The Effect of Shipboard Manning Levels on Operational Effectiveness - An LCS Case Study: Phase 3* (Final Report Summary No. NPS-N16-N210-A). Retrieved from Monterey, California. Naval Postgraduate School website: https://calhoun.nps.edu/handle/10945/57742

Steward, D. V. (1981). The design structure system: A method for managing the design of complex systems. *IEEE Transactions on Engineering Management*, *EM-28*(3), 71–74. https://doi.org/10.1109/TEM.1981.6448589

Sutcliffe, A., & Sawyer, P. (2013). Requirements elicitation: Towards the unknown unknowns. *2013 21st IEEE International Requirements Engineering Conference (RE)*, 92–104. https://doi.org/10.1109/RE.2013.6636709

Tokadlı, G., & Dorneich, M. C. (2018). Development of Design Requirements for a Cognitive Assistant in Space Missions Beyond Low Earth Orbit. *Journal of Cognitive Engineering and Decision Making*, *12*(2), 131–152. https://doi.org/10.1177/1555343417733159

United States Government Accountability Office. (2014). *Littoral Combat Ship: Deployment of USS Freedom Revealed Risks in Implementing Operational Concepts and Uncertain Costs* (No. GAO-14-447). Retrieved from https://www.gao.gov/assets/670/664672.pdf

United States Government Accountability Office. (2017). *Navy Force Structure: Actions Needed to Ensure Proper Size and Composition of Ship Crews* (Report to Congressional Committee No. GAO-17-413).

Wang, S., Lilienfeld, S. O., & Rochat, P. (2015). The Uncanny Valley: Existence and Explanations. *Review of General Psychology*, *19*(4), 393–407. https://doi.org/10.1037/gpr0000056

Woods, D. D. (2015). Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety*, *141*, 5–9. https://doi.org/10.1016/j.ress.2015.03.018

Appendix A. Wallpaper Submissions

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Lessons Learned																
What human data do we need to know and capture?	●						●								●	
How to balance the need for personal privacy vs. the need for human state sensing? E.g., Intra-vehicular activity (IVA) robots	●						●									●
We should display less data and show more knowledge because humans do not intuitively think statistically	●	●					●								●	
How to measure human-robot communication degradation over the course of a mission? Should the measure focus on human or robot?		●					●						●			
How to model after human adaptiveness to develop adaptive autonomous systems?		●					●							●		
How should an autonomous vehicle manage its own lessons learned process? Cf. ISS payload anomaly report system		●					●								●	
Capture knowledge from former/existing astronauts for "expert aids"			●				●							●		
Capture flown crew members' real-world expertise and make it available to in-flight crew			●					●						●		
Intelligent systems should provide rationales for their actions. This information helps to understand performance and optimize algorithms			●				●							●	●	
Flight software should not prevent real-time commands within a given range and values unless their execution results in hazards (lesson learned from 2013 ATCS flow control valve failure)				●			●								●	

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Use blind test data infrequently to avoid overfitting				●			●							●		
Data without a standardized meta-data model is dangerous						●	●							●		
All cognition is situated. What is the right context to capture to make trust assessment possible and correct?						●	●								●	
Simplified crew interfaces consistent across systems	●							●							●	
Operators typically need to know: communication failures, impact to system, workaround to operate system, time information		●	○				○	●							●	
Research “brain books” to learn more about Mission Control Center tools to help the crew		●						●							●	
Brain books research		●							●						●	
How to determine the priority/preference of tasks to be performed by the crew or vehicle autonomy?			●					●					●			
Check out SailorBob.com (forum for Surface Warfare Officers) for lessons learned				●				●								●
Many factors other than the quality of the solution affect its implementation (e.g., operator preference, constraints, conflicts with agendas)				●				●							●	
How to help users develop effective mental models of autonomous systems? Usable security research has shown that incorrect mental models can cause users to fall back on experience and expectations that are incorrect, causing detrimental consequences						●		●						●	○	
Autonomous systems that are capable of diagnosing themselves and displaying their malfunctioning states would be beneficial.						●		●						●	○	

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
This would prevent the crew from following faulty suggestions.																
Payload Operations Integration Center (POIC) has a near real-time system for retrieving real-time and recorded data for trending, troubleshooting and analysis	●								●					●		
Capability to verify that the crew have selected the most optimum solution		●							●					●		
Modeling needs to consider: average references, critical events, and black swans			●						●					●	●	
Autonomous systems work towards repeatability as input grows. Explain rationale if answer is different.				●					●						●	
There is a need of data integration to identify standards and guidelines	●									●					●	
Spacecraft specific "Watson" or "Siri" acting as onboard integrated electronic tech manual		●								●				●		
Systems should afford multiple levels of automation to adapt to crew needs			●							●					●	
High assurance for critical systems				●						●				●		
Automation can only compensate for loss of human capability (e.g., cognitive impairment) if all human tasks are known and can be programmed	●										●					●
Plan workload aiming for workload at the major malfunction	●											●			●	
Consistent design across systems to allow generalizable training/design for trainability				●								●			●	
Quality of systems integrated from parts made by different vendors should not be assumed						●					●					●
Need requirements to prioritize human-system design over strictly engineering requirements to ensure a system that human can use		●										●			●	

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Training for expertise with autonomy						●						●	●	●		
When building complex AI systems it can be very helpful to have separate teams build the AI and build the tools to evaluate the AI and diagnose failures. With one team the tools are always low priority						●						●				●
Technology																
Eye tracking to determine visual attention, situation awareness, and workload	●						●							●		
Facial recognition	●						●							●		
Need technology to convey uncertainty and trustworthiness so to calibrate trust				●			●							●		
What does monitoring complex systems for emergent behaviors look like for autonomous enabling technologies?	●							●						●		
Capability to conduct and evaluate onboard training of required fidelity				●				●						●		
Concept maps					●			●						●		
Assumption tracking					●			●						●		
Case-based reasoning systems					●			●						●		
Identify alternative to real-time internet to support crew recreation and well being						●		●						●		
Capability to project "what-if" scenarios in anticipation of less familiar or risky situations	●								●					●		
Lessons learned from NEXTStep simulations can feed into "risk of lessons" model		●							●					●		
Tableau (interactive data visualization)			●						●					●		
Transparency and trust remains design challenges. Links between understanding and authority/responsibility also needs analysis			●						●				●			
Resilience systems that reconfigure to accommodate lost functionality and support crew reconfiguration			●						●					●		

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Do certain AI methods afford more explainability than others?						●			●				●			●
Human guided machine learning			●							●				●		
How to formally verify the functionality of learning systems? How to avoid cyber-threats such as adversarial inputs that produce unanticipated outputs?				●						●				●		
How to orchestrate information flow among multiple intelligent agents (people or technology) with different capabilities to arrive at optimal solutions, according to time pressure, task configuration, etc?						●				●			●			
Jupyter notebook for integrating code, data, analysis, and visualization						●				●				●		
Compelling, easy to use, generate visualizations impact of changing resource management & utilization			●								●			●		
Using containers for applications to reduce underlying computing differences			●								●			●		
Smart technology needs to support flexible, dynamic coordination and distribution of work, probably with some constraints. Designing what flexibility can be anticipated, and what unexpected needs may arise is a different problem			●								●					●
Need to know which crew members are sufficiently trained to execute a procedure safely				●							●			●		
Design systems to emphasize communication, information transfer and situation awareness building. V&V of the information content to assess situation awareness.			●									●			●	
How to heal broken trust – human to technology, human to human			●									●	●			

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Need technology for automatically capturing lessons learned					●							●		●		
These AI systems will need lifelong learning (a la DARPA Microsystems Technology Office Lifelong Learning Machines project)					●							●		●		
How do we incorporate uncertainty (formally) into the models, estimates and processes? Especially given the little data we have *a priori* on these systems?					●							●		●		
Standards & Guidelines																
What attributes do we want to emphasize in standards? (e.g., simplicity)	●						●								●	
How much of crew schedule should be allocated to replicating ground functions?	●						●				●		●			
Crew and ground should be informed of fault detection status			●				●								●	
System shall promote share mental model among all agents-human-intelligent systems-robots etc						●	●								●	
Standards and guidelines on different modalities	●	●						●					●			
Standards and guidelines on integration of information		●						●							●	
Standards and guidelines on adaptation systems		●						●							●	
Ensure data collection in training and operation permit annotation and event marking			●					●							●	
Design guidelines for conversational & polite systems				●				●							●	
Standards and guidelines for transition to and from unscrewed vehicles				●					●						●	
Systems shall monitor crew procedure execution in real-time and provide				●					●						●	

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
feedback (cf. current flight control monitoring of crew procedure execution)																
Make sure test cases cover all of the components of functionality				●					●						●	
New technology transforms work. How to conduct task analysis for novel "under development" work settings?						●			●				●			
Automation must be flexible enough to support query types that programmers did not predict		●								●				●		
Data for logistics and spares must support Ops from end to end		●								●				●		
There shall be a capability for crew to create new procedures rapidly and efficiently			●							●				●		
Recommender systems should be preferred over single-selection "do-this" systems in cases where such a system improves trust			●							●					●	
All systems should eat their own outputs					●					●						●
Can human-autonomy relationship be better cultivated over time by having autonomy perform "sub-optimally" in non-critical moments so to facilitate the development of calibrated trust?						●				●			●			
Can systems be simplified or even the entire vehicle be simplified, to require a minimum number of tools/skills to operate and maintain?				●							●			●		
Standards and guidelines should provide a broader view of how automation, robot, and human activities can be flexibly coordinated beyond "levels of automation"						●					●					●
How to design systems with "hooks" that facilitate the integration of advanced intelligent capabilities?			●									●		●		

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Be able to create specialized display needed due to failures (or other reasons) and create info on extremely slowly changing failures			●									●				●
Make working with an AI or other virtual agent more like working with another person because people already know how to interact and work with people				●								●		●		
How to vet information from intelligent systems?				●								●		●		
How to improve/enforce HSI participation in Programs?					●							●				●
Learn from the design and construction area (architecture and engineering) about standard data models that are semantically rich such as the Industry Foundation Classes (IFC)						●						●				●
Out of Box																
How to communicate what the autonomy is thinking?				●			●							●		
Integration of smart environments to help crew make decisions	●							●						●	●	
How to visualize the historically separate flight and ground systems at the same time? How to overcome International Traffic in Arms Regulation (ITAR) and Export Control (EC) challenges?		●					●	●						●		●
How to identify the full complement of autonomous factors? Need to ensure levels and models of autonomous systems are addressed with criteria			●					●								●
Does culture play any role in determining whether crew or MCC get to manage the consumables on the ISS?				●				●								●

	Anomaly Response Steps						Resources and Processes						Relevance			
	Monitoring & Detection	Diagnosis	Solution Development & Evaluation	Solution Implementation & Verification	Resolution Documentation	Out of Box	Data	Knowledge	Analysis	Synthesis	Resource Management	Out of Box	Human Behavioral Research	Technology Development	HSI Considerations	Out of Box
Need to think of environments (hab, vehicle, EVA) as a system that can support users cognitively and physically by becoming autonomous						●		●							●	
Crew needs presentation of information, such as onboard vehicle design data (OVDD), most relevant to situation quickly to support diagnosis and solution formation		●							●					●		
How suitable are technologies based on machine learning, artificial intelligence to serve as information generator?					●					●						●
Intelligent systems need to be able to monitor/perceive people and learn/exploit models of them to communicate effectively	●										●				●	
HSI community needs to be stakeholder for comm improvements -- advocates for importance of comm			●								●					●
Drivers on crew autonomy: comm delay or loss, no escape to Earth						●					●					●
Engineered systems, training and operational policies need to be designed together but in reality are developed in silos due to different pieces being constructed by different entities. How to address this problem?						●					●					●
How to determine where best to allocate intelligent systems given limited budget and radiation challenges?	●											●				●
Need research on trust – how to address the autonomy valley	●											●				●
Change our model of designing for the masses toward crew customization	●											●				●
For on-orbit/in space failures, how do you autonomously determine understanding of failure is "good enough"? Sometimes root cause cannot be determined in vehicle		●										●				●

Appendix B: List of Technical Interchange Meeting Attendees

Name	*Affiliation*
Claudia Acemyan	NASA/KBRwyle
Bernard Adelstein	NASA-ARC
Brooke Allen	NASA
Keith Arthur	NASA
Immanuel Barshi	NASA Ames
Mario Berges	Carnegie Mellon University
Brent Beutter	Human Systems Integration Division, NASA Ames Research Center
Dorrit Billman	SJSU @ NASA Ames
Laura Bollweg	NASA - Johnson Space Center
James Broyan	NASA-JSC
Mark Chung	Verdigris
Mary Connors	NASA/ARC
Rick Davis	NASA
Donna Dempsey	NASA
Charles Dischinger	NESC
Terry Fong	NASA
Kelly Foreman	IBM
Jack Gale	NASA Ames
Javier Garcia	US Army Research Laboratory
Brian Gore	NASA Ames
Matt Guibert	NASA Ames
Angie Haddock	NASA
Steven Hillenius	NASA
Jon Holbrook	NASA LaRC
Kritina Holden	Leidos at JSC
Jeffrey Homola	NASA Ames Research Center
Sandy Howard	ARL
Matthew Johnson	IHMC
Edward Graham Katz	IBM
Thomas Keeping	NASA
Jeff Kephart	IBM Research
Gary Knickerbocker	MSFC
Michael Krihak	USRA/NASA ARC
Kara Latorella	NASA Langley
Patrick Lincoln	SRI
Victor Luo	NASA JPL
Jessica Marquez	NASA Ames
Kaleb McDowell	Army Research Laboratory
Mark McElyea	NASA - MSFC
Kerry McGuire	NASA
Andrew Mishkin	JPL

Name	*Affiliation*
James Murdock	IBM

Cynthia Null .... NASA Engineering and Safety Center
William Paloski .... NASA - Johnson Space Center
Karen Pickering .... NASA - Johnson Space Center
Stephen Robinson .... UC Davis
Natalia Russi-Vigoya .... KBR Wyle
Brandon Schmitt .... NASA Ames
William Scroggins .... U.S. Army Research Institute
Matt Sharpe .... NASA Ames
Nita Shattuck .... US Navy Naval Postgraduate School
Jay Shively .... NASA - Ames Research Center
Alonso Vera .... NASA Ames Research Center
Chengtao Wen .... Siemens Corporate Technology
Elizabeth Wenzel .... NASA Ames / Human-Systems Integration
Alexandra Whitmire .... HFBP Element
Tom Williams .... NASA/JSC
Reed Williams .... Siemens CT
Shu-Chieh Wu .... NASA Ames / SJSU