Implementing Artificial Thinking Autonomy with Model-Based Systems Engineering

Mitchell Kirshner
Steward Observatory
933 N Cherry Avenue
Tucson, AZ 85705
mkirshner@arizona.edu

Fernando Figueroa
NASA Stennis Space Center
Autonomous Systems Laboratory
Stennis Space Center, MS, 39529
Fernando.Figueroa@nasa.gov

Lauren W. Underwood
NASA Stennis Space Center
Autonomous Systems Laboratory
Stennis Space Center, MS, 39529
Lauren.w.underwood@nasa.gov

Abstract— **Complex autonomous systems capable of successfully operating independently under 'known unknowns' and harsh conditions require paradigm innovation in modern development strategies. In the field of autonomy, developing a system-of-systems which can ostensibly think for itself in the face of 'unknown unknowns' is still a field of ongoing research. Maturing the systems architecting and modeling methodologies for developing henceforth named Thinking Autonomous Systems, which are verified with digital mission simulation, can potentially usher in the next generation of artificial intelligence for space exploration.**

The concept presented in this paper incorporates multiple Model-Based Systems Engineering and simulation methodologies combined as a new paradigm to design a novel, biomimetic thinking autonomy strategy. Anachronistic concepts from classical Kantian philosophy will be leveraged to inspire architectural designs that could be used for complex distributed systems in deep space. To accomplish this, digital transformation of a document-based implementation plan for Thinking Autonomous Systems, generated by experienced NASA software engineers, is implemented for NASA's Platform for Autonomous Systems by creating descriptive and executable software models in SysML to prototype real-time operating capabilities.

This conceptual implementation has been developed by incorporating model-based digital simulations to theorize how a cyberphysical thinking system would achieve specific strategies without crew reliance, while simultaneously being resilient to all operating conditions and remaining functional when devoid of ground communication. Additionally, ensuring that an autonomous system framework is an ethical Artificial Intelligence requires careful consideration of system behavior and accountability, human factors for teaming with a thinking autonomous system, and comparison to other modern approaches used for implementing true autonomy.

This paper presents the first steps in formalizing the metacognition required for instantiating a truly Thinking Autonomous System; the approach described symphonizes autonomy characteristics from classical philosophical into a unified software architecture describing human thought. In the future, the foundational models described in this paper can be further leveraged to help advance research into thinking autonomy requirements for future deep space missions as well as for current near-term applications, i.e., living aboard crewed spacecraft like a NASA Gateway cislunar habitat.

TABLE OF CONTENTS

1. INTRODUCTION 1
2. LITERATURE REVIEW 2
3. METHODOLOGY 4
4. RESULTS 6
5. CONCLUSION 8
REFERENCES 10
BIOGRAPHY 10

1. INTRODUCTION

Autonomy, as defined by NASA, is "the ability of a system to achieve goals while operating independently of external control" [1]. Historically, 'brute force autonomy', which describes algorithmic processing of all possible circumstances before decision making based on strategies developed by experts offline, has been traditionally used to implement autonomy. However, as systems become increasingly complex, this method, which requires accounting for all "what-if" scenarios, becomes nearly impractical to both implement and be completely comprehensive, since inevitably, there will be cases which arise that experts overlooked [2]. Additionally, brute force autonomy strategies require continual maintenance and updates, which eventually becomes problematic for implementing with "real-time" systems that will be remote/inaccessible, like those planned for the NASA Artemis Gateway cislunar space habitats, Mars, deep space missions, and beyond. Furthermore, these methods restrict generalizability, requiring engineers to continuously develop brute force strategies *ad hoc*.

Alternatively, enabling autonomy with a novel strategy, referred to as 'thinking autonomy', which is ability of a cyberphysical system (CPS) to implement a biomimetic (mimicking nature) human thought process architecture, is an approach that could be used to address some of the problems and limitations associated with the current brute force implementation methods described above. Thinking autonomy requires both a method of self-awareness for a system to understand its own operating conditions, as well as the ability to model behaviors for self-directedness to achieve mission objectives and goals [1]. Thinking Autonomous Systems (TAS) also exhibit self-sufficiency and the ability to

979-8-3503-5597-0/25/$31.00

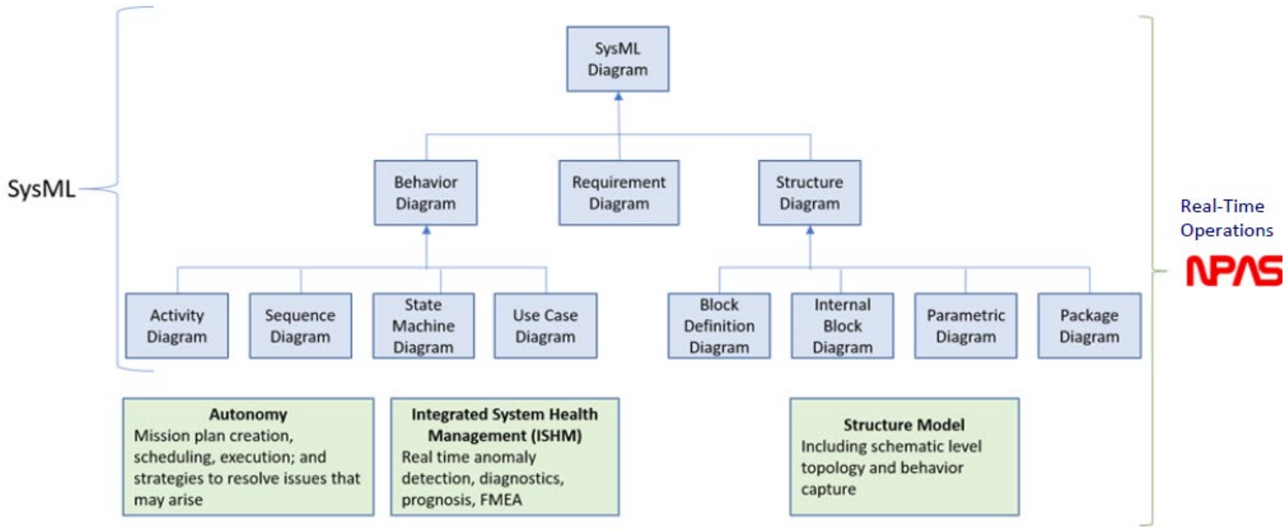


Figure 1. NPAS comparison to SysML for MBSE, Real-Time Operations, and Autonomy.

operate independently, without external support. Irrespective of operating conditions, a TAS must be resilient enough to address unforeseen events encountered that could interfere with the system's ability to safely achieve mission goals and objectives.

Additionally, a TAS must encompass knowledge and understanding of the system (itself) within the context of a hierarchical network comprised of other TAS systems: for example, a constellation of autonomous satellites [1]. New paradigms for implementing system autonomy would greatly benefit, for example, US government national goals for implementing Position, Navigation and Timing (PNT) in cislunar space, as defined by the White House Office of Science and Technology Policy's Cislunar Science & Technology Strategy [3], [4]. Therefore, to effectively develop a software TAS, there is a need for specialized software that can manage real-time operations, simulations, and asynchronous operations, all of which are typical aspects of a complex CPS.

To begin addressing the challenges associated with developing a TAS, Stennis Space Center's Autonomous Systems Laboratory (ASL) has created an innovative software tool, the NASA Platform for Autonomous Systems (NPAS). This platform provides direct solutions that help mitigate the drawbacks associated with current modeling languages in architecting and simulating a TAS. As shown in Figure 1, the drawbacks of SysML, a commonly used systems architecting language, include: a cumbersome simulation interface, a lack of native support for integrating domain-specific digital design elements, and a limited selection of visualization for systems analyses [3]. NPAS enables all capabilities facilitated by SysML and simultaneously has the ability to expand these tools even further by adding the capability for real-time operations, autonomous strategy implementation, and integrated system health management [2].

2. LITERATURE REVIEW

Perhaps the largest wealth of compiled knowledge associated with processes that encompass human thinking and autonomy is derived from the work of classical philosophers like Immanuel Kant, Georg Hegel, and Arthur Schopenhauer. These philosophers have each compiled manuscripts that describe their interpretation of the metaphysical decomposition of different systems associated with human thought; these analyses are still used and studied to this day. Then, in 2013, a book was written by former JPL project manager and chief of Voyager general science data, Szabolcs de Gyurky, that established a document-based architectural framework for a software TAS [5]. This work coalesced information from these past great philosophers which then in turn, is used to synthesize a technical baseline for developing next generation "thinking" systems.

In the following sections of this paper, this interpretation of creating a theoretical basis for human thinking as it applies to implementation of thinking autonomy is summarized, as well as analogous applications for existing deep-space mission strategies, to describe how the use of autonomous navigation strategies within the framework of the theorized TAS could be implemented. In theory, this capability would then be able to facilitate high-fidelity simulation of a TAS that could be used to make decisions, based on the system's ability to "think" without human intervention, i.e., in mission settings, which would include, for example, remote environments such as cislunar space. The system's ability to "think" would in turn have the potential to increase the technology readiness level of mission solutions and correspondingly help increase the safety of astronauts, who may, for example, be put in situations whereby instantaneous spacecraft maneuver decisions need to be made to avoid hazards like interception with space debris or another spacecraft.

The Autonomous System: A Foundational Synthesis of the Sciences of the Mind— In this document-based architecture for TAS presented by de Gyurky and Tarbell [5], the authors put forth the proposition that key information about the hierarchical cognition systems and subsystems of human thinking and autonomy has to be properly defined to effectively model and simulate interactions which would formalize through software human metacognition and autonomy. de Gyurky and Tarbell sought to formalize these definitions through document-based systems engineering, by coalescing the processes of thought as defined by classical philosophers like Kant, Schopenhauer, and Hegel.

According to de Gyurky and Tarbell's summarization of these classical philosophers' interpretation of thought, there are eight underlying systems which facilitate inter-system operations: Will, Reason, Intellect, Sensory, Understanding, Presentation, Decision, and most importantly a central Thought system. These systems are graphically depicted in Figure 2, as related to their roles in internal and external communication for operations/functionality of what is being defined as the Thinking Autonomous Systems (TAS).

In theory, these systems mimic the processes that form the foundation of human thinking. Then, to describe how these systems communicate to enable human thinking, de Gyurky takes this interpretation one step further, and provides a representative analogy as follows: "Take as an example an object like a hand grenade, which the Will system wants to physically pick up, but does not know what it is. So, the Intellect system adds to the object being viewed, 'That is a hand grenade; it is dangerous to our existence.' The Understanding system provides the Form of All Possibilities and adds, 'If we pick it up, it could be useful; but if it is booby-trapped, it will destroy us'" [5]. This descriptive and rich scenario provides an example of how thinking incorporates multiple processes and associated systems of thought, all of which would be necessary to implement TAS operations and demonstrates how this capability would be required for instantaneous prioritization of tasks to ensure human safety and survivability.

Existing Autonomy Strategies—NASA Stennis Space Center's Autonomous Systems Laboratory has demonstrated numerous autonomy applications using NPAS (built upon the Gensym G2 software environment) that incorporates: Integrated System Health Management (ISHM), domain model development (digital twin), mission creation and management, distributed autonomy, auditory Human-Computer Interaction, and network bridges [6], [7], [8], [9], [10], [11]. Application domain areas include electrical, fluid, mechanical, communication and network systems; the platform is architected in a way that quickly and readily enables other systems (e.g., human interaction systems) to be added, as needed. NPAS facilitates a software mechanism to implement real-time thinking autonomy that can leverage the baseline framework that is created using SysML. In this way, NPAS can instantiate a TAS as identified in Figure 2 as a form of artificial intelligence.

However, as previously mentioned, the way artificial intelligent systems for autonomy are currently developed involves using an approach that can be referred to as "brute force." Often, brute force autonomy strategies, including those implemented through neural networks and machine learning, are custom-tailored for very specific problems or

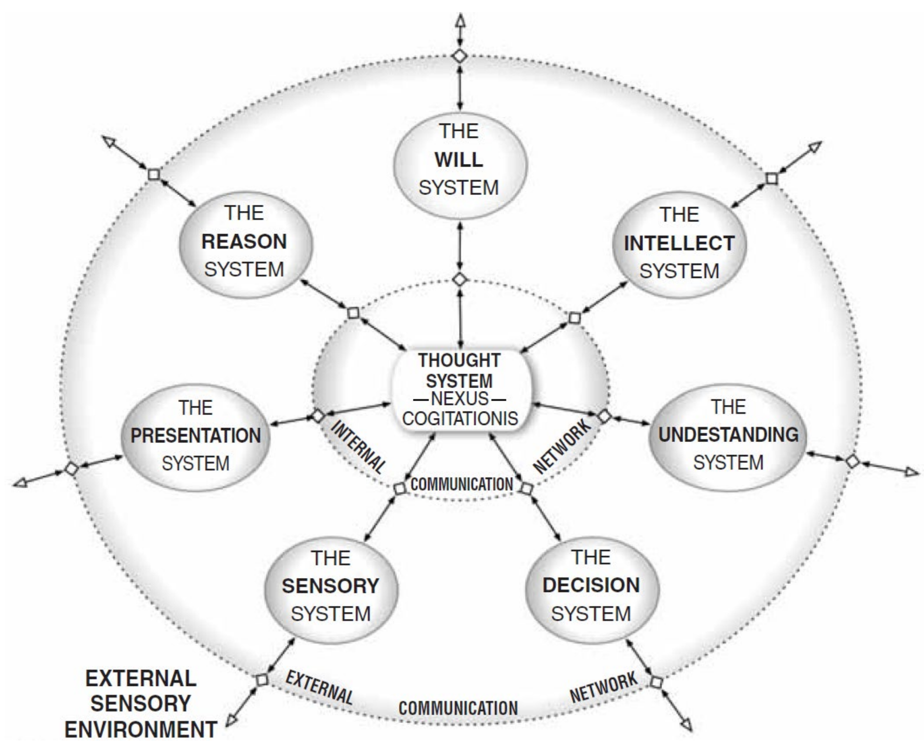


Figure 2. Thinking Autonomous System architecture based on classical philosophy [5].

use-cases. Therefore, reusability and evolution of these implementations becomes problematic and/or extremely difficult to achieve; autonomy strategies in one domain space cannot always be transferred to other science and/or engineering applications [12], [13], [14], [15], [16], [17]. Additionally, autonomy implementations may include, for example, cloud computing, robot swarm coordination, medical diagnostics, naval or maritime situational awareness, and resource management. For the purposes of this paper, the focus will be on autonomy for deep space and cislunar mission applications, which then could be leveraged for multi-mission capabilities.

Autonomy for Cislunar Position, Navigation, and Timing (PNT)—Referring back to the proposed example presented in this paper for implementing a TAS, PNT: currently, there are technology gaps associated with implementing PNT autonomy in cislunar space. This is because most traditional methods of PNT in space rely upon a spacecraft's proximity to Earth. However, in cislunar space, a spacecraft would have to wait several seconds to receive signals from those Earth-based positioning systems; this delay in data transmission receipt, in conjunction with the additional uncertainty that the signal may not be received at all, as well as the fact that this wait time for signal receipt could delay and amplify the uncertainty in identifying positions, all collectively could impact the spacecrafts ability to timely implement effective collision avoidance maneuvers. This circumstance arises because in cislunar space, spacecraft distances will be too far from positioning data providers, like Earth's Global Navigation Satellite Systems. Therefore, this data gap introduces technological hurdles that must be overcome when implementing, for example, lunar habitats, since access to accurate satellite location coordinates through PNT for safe and trusted rendezvous and proximity operations (RPO) is not currently available in cislunar space [18].

Therefore, implementing operational autonomy far from Earth, for example for cislunar PNT, remains in its nascent stages of development, with papers only just being published within the last two years that define and parametrize the architectural design space [19]. There are other studies that assess specific solution feasibility, like using the timing of X-Ray pulsar flashing to conduct 'XNAV' navigation in deep space. And still other studies assess GNSS-type systems placed on the moon; however, these solutions are not transferrable to missions operating in deep space because proximity to Earth's moon is required [20], [21]. Fortunately, studies do exist that describe visual-based PNT solutions similar to XNAV using sidereal tracking and mutual navigation: celestial navigation, inspired by Polynesian wayfinding [22], [23]. These studies provide a basis for developing deep space navigation and autonomy using satellite constellations and timed measurements of starfields.

3. METHODOLOGY

To develop a TAS natively capable of operating on-board a spacecraft, the intent here is to use Model-Based Systems Engineering (MBSE) to plan and architect a TAS first, prior to software porting or interoperability with the autonomy platform, NPAS. MBSE requires a modeling language, a methodology for that language, and an authoring tool: for this study's system architecting endeavor to define the architectural hierarchy of a TAS for autonomy and PNT, the language SysML, its associated Object-Oriented System Engineering Methodology (OOSEM), and Cameo Systems Modeler authoring software were used.

Cameo Systems Modeler's Simulation Toolkit plugin allows execution of SysML diagrams, which are forms of code abstracted through visual programming languages. Through the capabilities afforded by this simulation toolkit (Cameo), SysML models can connect to external software, for example NPAS, using digital thread, forming the foundation of a TAS digital twin [24]. Then, with the knowledge that NPAS provides a suitable toolset to model a TAS with real-time asynchronous communications between them, system engineers can manifest the type of thinking system described by de Gyurky and Tarbell [25].

Therefore, to attempt to construct a truly "thinking" autonomous system, engineers and architects must approach understanding the TAS architectural framework from all prospective levels, including a top, mid, and lower-level viewpoint to capture the hierarchical functions to code processing systems for thinking and system communication. Table 1 defines a summary of the descriptions for the top-level systems in the proposed "thinking" system-of-systems that provides an overview of each of respective system's responsibilities and is represented here as a compiled synopsis of published literature [26].

Figure 3 represents a hierarchical structural decomposition, showing subsystems of a TAS, and includes some lower-level descriptive components to help illustrate the next level of complexity that is involved with this system-of-systems hierarchy. In this figurative subsystem representation, systems are grouped according to likelihood of interoperability with one another; each system also identifies the philosopher whose work can be attributed to the underlying rationale behind this biomimetic strategy for implementing "thinking" autonomy. Each subsystem will still require further development to determine respective representation within a CPS.

For TAS implementation, this theoretical 'thinking" system will predominantly rely on leveraging capabilities enabled by NPAS, which is comprised of knowledge bases - containing rules and laws for the Reason system, logic for handling abstract concepts in Intellect system and operations in the Will system, as well as leveraging the field of computational ontology, which can bridge knowledge across systems, to establish 'intra-system' system communications. The field of

Table 1. Top-level description of subsystems in a Thinking Autonomous System.

System	Description
Will	The "CEO" of the mind; any system of thought must have a Will. Imparts the mandate to operate with the mind.
Reason	Advisor to the Will; constantly validates the Will's intended actions through processing rules and axioms both hard-coded and learned.
Intellect	Enables the processes of ingesting, integrating, and categorizing all information in the constellation system; the 'librarian' of the mind.
Sensory	The entry point to the mind. Receives external stimuli called phenomena and provides 'environmental awareness' of both internal and external state, giving rise to self/non-self-discrimination necessary for self-awareness.
Understanding	Identifies, classifies, and distinguishes phenomena for use by Intellect and Decision to prepare decision-making by Will. Leads to the capacity to acquire and apply knowledge.
Presentation	The 'theater' of the mind; presents to other systems the consequences of any action under consideration by Will.
Decision	Provides a constant stream of reasoned decisions for Will to consider. Assembles data, activity, communications, and products from other systems into reasonable courses of actions for Will to take.
Thought	The nexus of all systems in the constellation of the mind. Monitors, assesses, manages, regulates, constellation activities with the goal of generating biomimetic, human like thoughts.

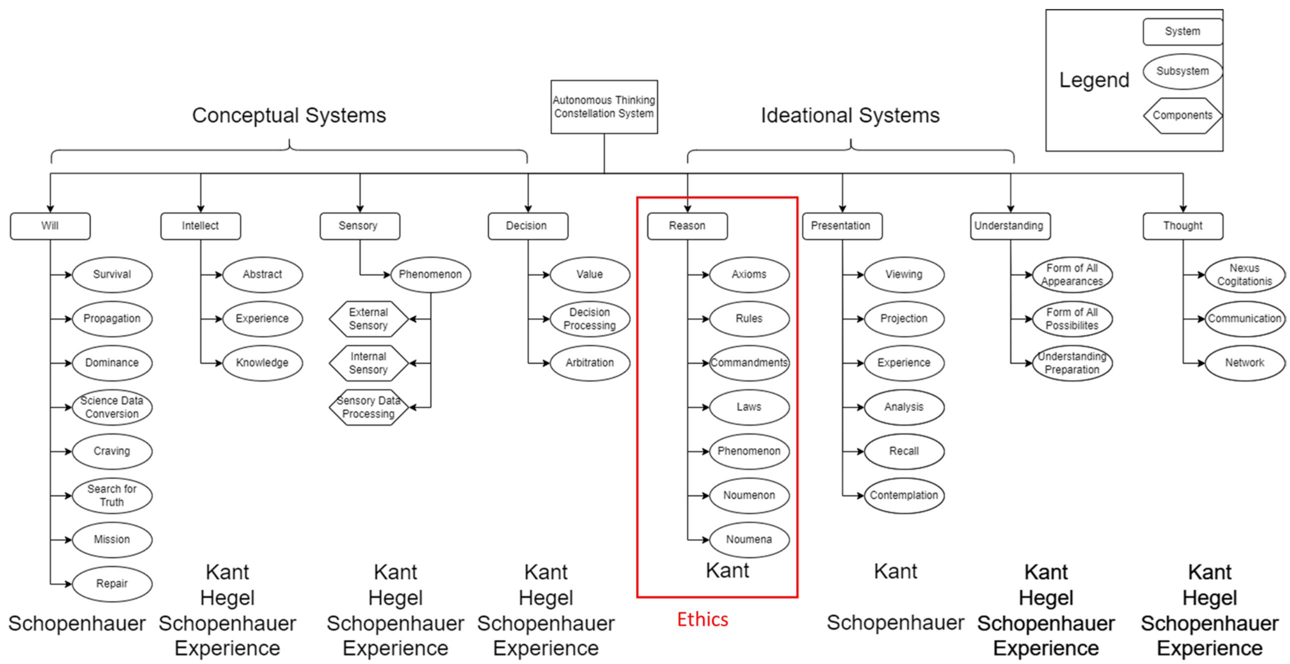


Figure 3. Diagram of classical autonomous subsystems, with sourcing from philosophers or engineer experience.

ontology provides the foundational definition of objects, as well as corresponding interrelations, and formally represents knowledge components of a domain or knowledge models [27]. Furthermore, since a TAS should constantly evolve and learn, ontologies should be dynamic and change over time: a thinking system should be able to change what it 'knows' as it receives or generates new information. From an ontological perspective, NPAS can simulate the modeled system-of-interest (the mind) in its respective environment of interest through knowledge bases [28].

Notionally, the TAS is architected in a way that takes into consideration performance metrics for cislunar PNT and PNT beyond Geostationary Earth Orbit (xGEO). Though a fully functional TAS implementation is beyond the scope of this paper, once a functioning technical baseline is developed, PNT performance could be used as a foundational parameter, which in turn, could be used to compare against conventional PNT technical performance measures and key performance parameters. Furthermore, some metrics, aside from Size, Weight, and Power considerations (SWaP), for assessing PNT performance that could also be integrated as TAS parameters include: positioning precision in units of distance within acceptable statistical error, data transmission speed, areal coverage of the PNT network, and portability to different operating environments (e.g. different cislunar orbits, low lunar orbit, asteroids, Mars, etc.) [21], [29], [30], [31], [32].

Then, a digital space mission engineering software ANSYS Systems Tool Kit (STK) and its associated Orbit Determination Toolkit could provide the computational toolset to rapidly assess the performance of the proposed TAS SysML models developed through, for example, using ANSYS ModelCenter digital thread integration, which enables the use of physics-based simulations for high-fidelity performance analysis for the proposed baseline thinking system [3], [24], [33]. This digital simulation approach would demonstrate empirical validation that the proposed thinking system autonomy strategy for PNT could perform similarly to other strategies in the context of present-day case studies.

4. RESULTS

The intent proposed in this paper is to have initiated model-based engineering of systems technology which biomimics the human mind, by combining all the information required to enable a TAS, synthesized from the treatises of Kant, Hegel and Schopenhauer and presented by de Gyurky and Tarbell. The initial steps for development include accumulating information resulting from research into techniques used for implementing modern and classical autonomous cislunar PNT system, which in turn facilitated the first steps in the digital transformation from document-based descriptions of a software TAS to a formalized MBSE platform [34]. This proposed technique enables model reusability and data sharing that can, in turn, be used to architect system-of-systems that can integrate operations that would become magnitudes too complex to implement with traditional brute force autonomy approaches and/or without human intervention [10], [35], [36]. Figure 4 depicts a SysML Block Definition Diagram (BDD) which represents the proposed structure for a TAS within this operational context.

In Figure 4, the systems in Figure 2 exist as SysML 'part properties' over the overall system-of-interest. Each of these part properties is 'typed' by a Block, like the ones describing the concept of CPS, NPAS, and the System of Interest (SOI). The stick figure Block represents a SysML 'Actor,' an atomic unit within SysML that cannot have part properties. The interaction between these part properties is summarized in Figure 5, which is an internal block diagram (IBD) of the system-of-interest Block. An IBD provides a 'White Box' view of the 'Black Box' Blocks in a BDD. The Figure 5 IBD combines the structural information of Figures 2 and 3 in a method accessible to programming languages.

For a hierarchical low-level view of the software components in a TAS, Figure 6 describes the proposed interpretation of communication modes described for the Sensory system in a TAS, again based on Kantian terminology. The small square port linking a connector to the outside of the system border in Figure 6 is the same port as in Figure 5: an example of the overall model viewed differently through different cross-cutting diagrams and context viewpoints. While Figure 5 does contain some 'White Box' elements, details are limited and reserved for graphical simplicity; Figure 6 shows greater TAS system fidelity in the SysML model, meaning that a deeper hierarchy is defined for subsystem and software program sets. SysML and OOSEM create hierarchy within systems by defining composition of blocks as a grouping of parts and references. This provides a basis for modeling TAS structure and component interactions such that several design alternatives for intra-system communications can be rapidly assessed.

The top left of Figure 6 shows a part of the TAS sensory system: the autonomous system state subsystem. For the representative example implementation used for this paper, this subsystem would house operations and algorithms that facilitate PNT methods of low computational overhead. With the digital transformation of classical metacognitive philosophy into a new framework for autonomy, xGEO PNT can now be described in a Kantian Framework. The following paragraph is an example of how this TAS could be instantiated from this descriptive framework (described above), illustrated as a novel TAS spacecraft implementation:

First, the system's Will system creates a requirement for the spacecraft to know where in its environment it is located . The Intellect system then draws from experience, identifying locations of nearby waypoints like planets, stars, or other recurring deep space objects as reference. The overall TAS tasks the Sensory system to provide additional data to the Reason system by communicating through the central *Nexus Cogitationis* in the Thought system. Then, the Understanding system assimilates all the synthesized information, creates a representation of the situational reality; the Presentation system simplifies this information for ingestion into the Decision system, which finally actuates subsystems to adjust the system as necessary to achieve the desired position and orientation to accomplish autonomous PNT.

The sequence for a TAS conducting PNT described in the previous paragraph can be modeled and simulated by extending the NPAS SysML structural model with cross-cutting sequence, activity, state machine, and parametric SysML diagrams. Then, upon modeling completion, digital transformation of this PNT TAS sequence could be tested and simulated in ANSYS STK as the subject of future research.

The proposed concept explored in this paper examines a methodology for comparing thinking autonomy to brute force autonomy, using PNT as a test case. In the future, this test case could be leveraged for assessing utility in, for example, a NASA Gateway system that would be required to use PNT to determine its own location for rendezvous and proximity operations with an incoming spacecraft/vehicle [38], [39]. To date, existing deep space PNT methods typically use brute force calculations during Initial Orbit Determination (IOD) [40]; in future iterations of the TAS software, PNT performance could be numerically compared directly to a brute force pulsar-based navigation architecture for the NASA Gateway, and assessed as a potential means for developing true thinking autonomy [41].

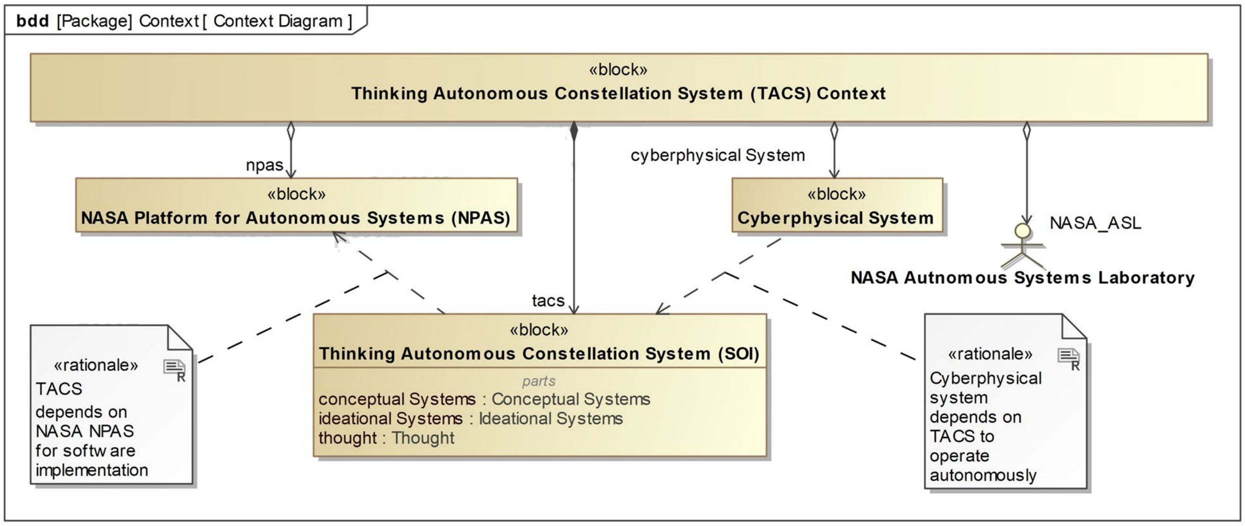


Figure 4. TAS in context with NPAS and its CPS host.

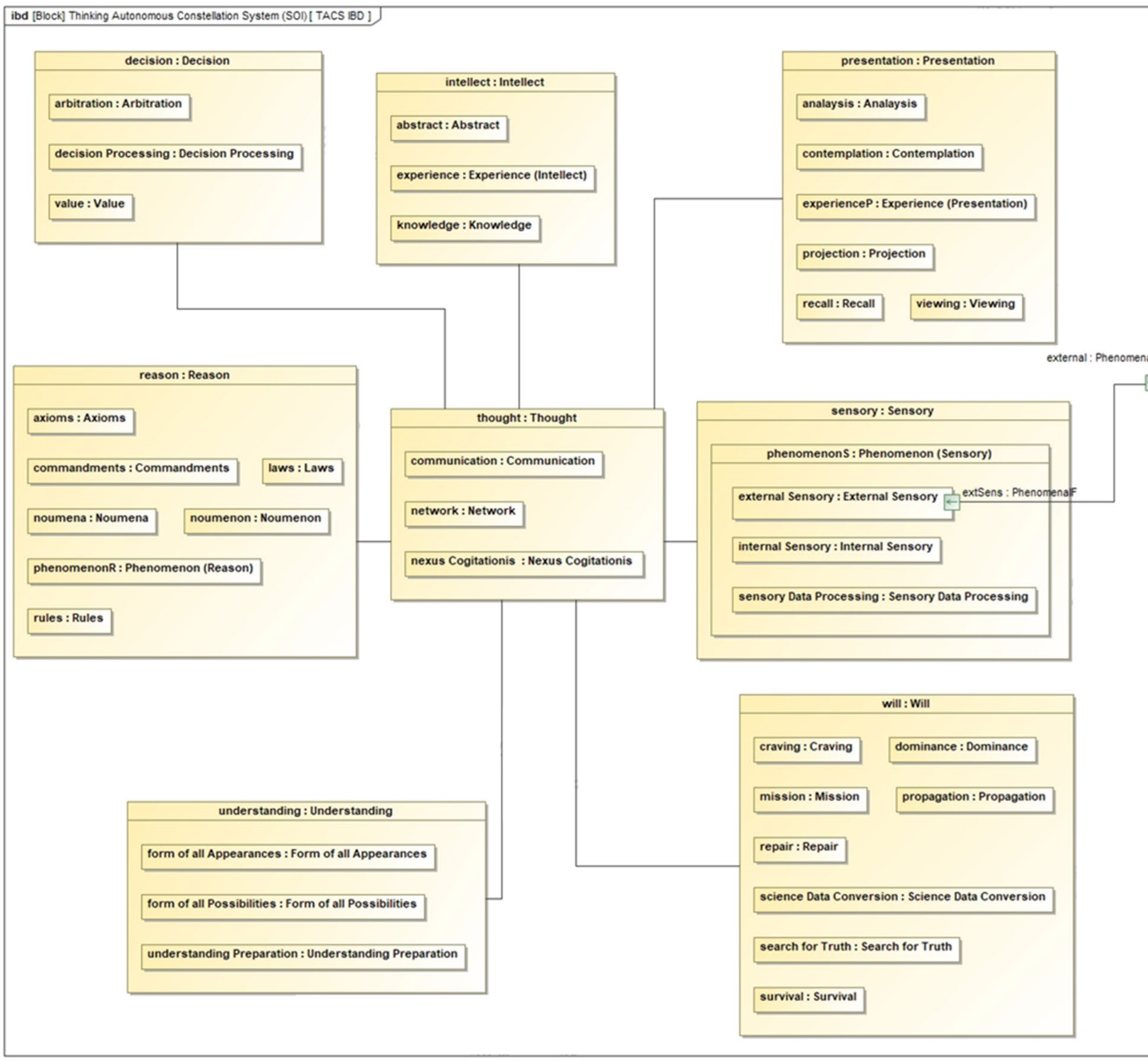


Figure 5. SysML IBD; digital transformation of Figure 2.

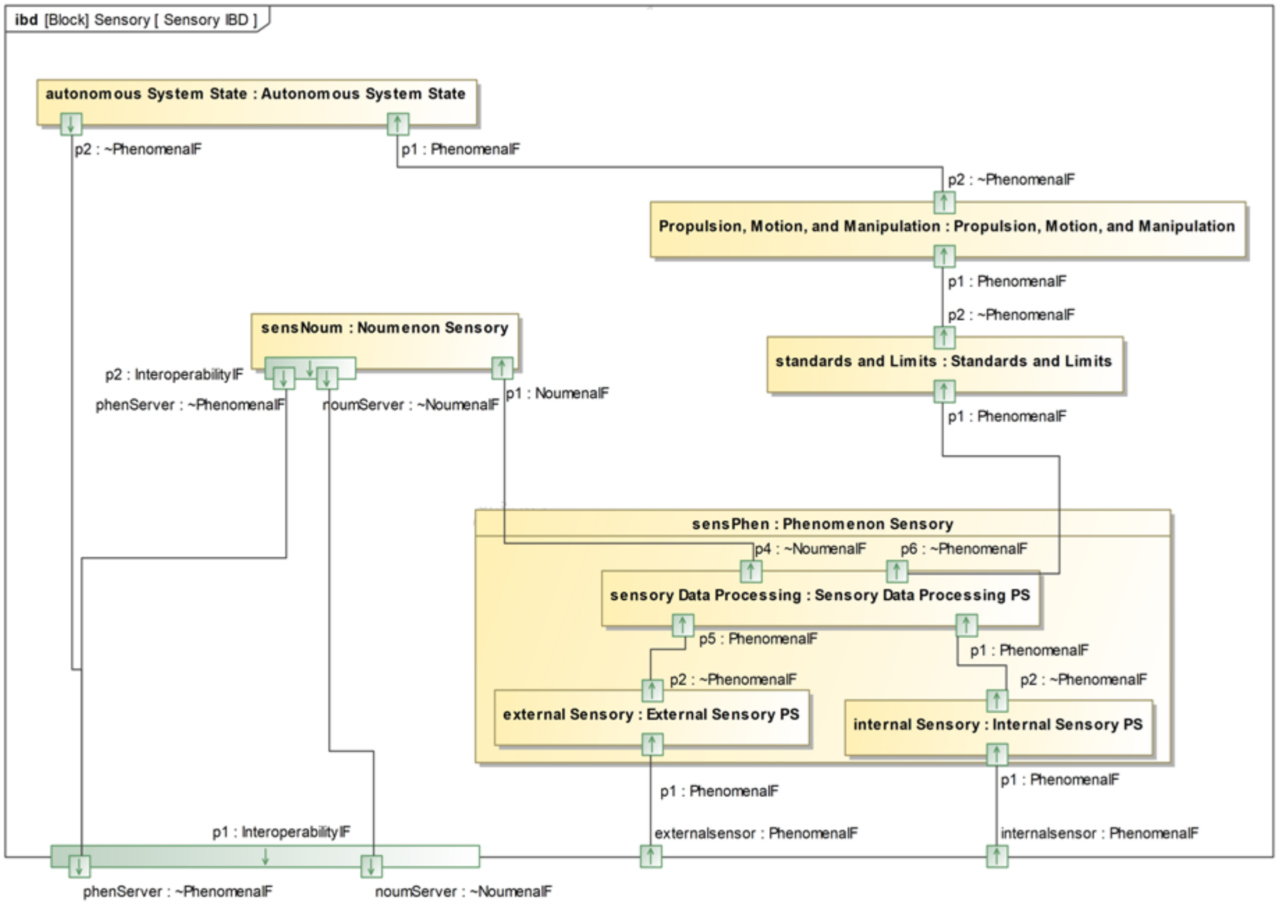


Figure 6. SysML low-level IBD for the Sensory system.

Furthermore, because thinking autonomy can in some ways be compared to Artificial General Intelligence (AGI), the development methodology presented in this paper also recommends future testing against benchmarks set for contemporary large language models (LLMs) and other current approaches that are used to mimic human behavior. One such benchmark that has been considered as a performance metric for the proposed TAS discussed in this paper is the Abstract and Reasoning Corpus for Artificial General Intelligence (ARC-AGI), a tool commonly used by tech companies like Google and OpenAI to measure efficiency of their developed tools [42]. Future iterations of the proposed TAS could focus on adapting its NPAS implementation to satisfy the computational requirements put forth by ARC-AGI that benchmark AGI performance and measure the humanness of the AI intelligence by proxy.

In this notional ARC-AGI use case, the TAS developed would diverge from the proposed PNT use case for NASA's Gateway cislunar space habitat. and transition towards an implementation that requires addressing ethics; in this way, the TAS development lifecycle methodology would benefit from both incorporating ethical considerations associated with the potential transitioning of AI use cases, and adherence to best practices in developing unbiased and safe autonomy solutions [43]. Source material for ethical artificial intelligence design guidelines could stem from a range of sources, including government and academic research [43]. Additionally, computer science industry leaders produce ethics reports which could also guide TAS development; for example, IEEE's Ethically Aligned Design provides information for developing Autonomous and Intelligent Systems [44]. These types of documents could supplement guidelines created by the United States and international intelligence communities for better understanding AI ethics and the implications of using autonomy strategies designed for a specific intended purpose in different domain areas [45], [46].

5. CONCLUSION

This work presented in this paper has delved into the foundational architectures required for implementing true thinking autonomy, which is required for enabling further expansion into novel strategies for complex cyber-mechanical systems to act autonomously in known/unknown scenarios, while simultaneously being resilient to all operating conditions as well as potential lack of external communication. Developing an autonomous CPS with the advanced cognitive skills of a human being is by no means a simple task; research presented here in this paper has only provided the preliminary foundation for advancing the development of a paradigm for architecting these thinking autonomous systems.

The theory presented directly lends itself towards advancing engineering capabilities required to develop, for example, an autonomous PNT system that meets NASA Artemis cislunar space habitat mission notional autonomy needs. However, implementing a fully functional TAS constellation using any tool, including NPAS, requires additional research beyond the scope of this paper. Furthermore, ethical scrutiny to guide engineering capabilities and ensure appropriate system use cases must also be considered when developing a software system that is intended to function fully autonomously, like the human mind. Based on compiled research to date, and the theory presented in this paper, ethical behavioral rules are housed within the Reason subsystem [47].

Due to the current size and resolution of the SysML model created for this research, traditional document-based reporting precludes the inclusion of all diagrams in their entirety; however, using Cameo Systems Modeler allows the ability to zoom and scroll through the larger diagrams of interactions in the cognitive systems. Little if any information postulated by de Gyurky and Tarbell mentions interactions of multiple TAS in swarm/constellation, and/or in hierarchical structure. These types of communications would be characteristic of operations in an autonomous constellation designed to complete multiple missions. Conceptualizing this software equivalent of 'mind-to-mind' system communication remains a subject of future research.

In theory, future scaling possibilities of, for example, thinking autonomy implemented across multiple satellite constellations in xGEO in different swarm configurations, could be made possible by the scalability functionality afforded by NPAS [37], [48]. Additionally, future testing on the types of novel autonomy strategies presented in this paper, using high-fidelity digital simulation, may require large computing power to test performance against conventional PNT methods; then, once a TAS is tested, verified, and validated as a functioning system, the goal implementation would be for a system like NASA's proposed Gateway space habitat to require less brute force computational power for on-board processing when conducting IOD [49].

This autonomy paradigm is not meant to completely supplant traditional methods of machine learning based on training data and classification; instead, it seeks to include them collectively within an overall framework that reduces the overall computational overhead of an onboard system: a necessary feature in remote, uncommunicable environments. By using digital thread technology to connect real-time NPAS models to system data sources, the intent is to create an algorithm repository that can be used to conduct trade studies of PNT solutions in cislunar space using simulation software like ANSYS STK [24], [33]. Furthermore, other important goals for any TAS use case include, but are not limited to, abiding by ethical AI standards, cybersecurity standards, and overall sustainability and equality goals of the international space community [50], [51]. Additionally, other areas of future investigation include:

- Low-level software component architecture.
- Human-computer interaction between TAS and crew.
- Trust metrics for Human Autonomous Teaming.
- Comprehensive ethical AI analysis on the potential use TAS for unintended purposes beyond space autonomy.
- Integration of SysML/STK software on NPAS.
- Developing generalized thinking autonomy strategies.

The longer-term goal is to develop a comprehensive digital basis for a trade study to optimize the architecture of a TAS that would be utilized for deep space operations. As the field of MBSE grows and SysML version 2 and its core language KerML becomes more widely available to users, additional capabilities for architecting autonomy may arise; however, to date, NPAS has preliminarily demonstrated the ability to afford the capability to support on-board real-time autonomy and scalability, and therefore has a greater technological and space-readiness advantage over modern SysML.

An additional noteworthy concept related to implementation of a TAS is the psychophysical action of implementing theoretical thinking autonomy in a real-world robotic example; this application creates the potential for new cybersecurity vulnerabilities that might not emerge or become evident until a fully realized TAS is functional in an operational capacity. In an effort to mitigate potential system security issues introduced by using a lower technological readiness system in a cybersecurity environment, heuristics compiled from several international space system standards, as well as the US Department of Defense System Engineering Research Center, can be leveraged. [50], [52], [53], [54]. Digital transformation of these heuristics leads to the development of requirements and corresponding use cases that define verification schema for system security [43]. The SysML and OOSEM methodology proposed in this paper, in combination with the notional TAS introduced, provides a platform to afford the formalization of cybersecurity requirements, and, in conjunction with the Cameo Systems Modeler Simulation Toolkit, would enable a means to automate verification to demonstrate that the TAS has met those requirements in relevant operating scenarios.

While not necessarily obvious to engineers without an education in classics, philosophy and the science of human nature has much to offer for architects seeking to develop system and system-of-systems autonomy. Historical concepts like classical philosophy and wayfinding can inspire solutions and design, leading to the development of functional systems which have the potential to revolutionize the field of thinking autonomy for complex constellation systems, as well as the field of artificial general intelligence.

The theories of the mind postulated by classical philosophers are difficult to test empirically; yet, emergent from this research is a novel schema by which to compare and contrast different architectures defining thinking autonomy from historical or contemporary sources. Designs for TAS swarms working in concert must transcend classical metacognitive philosophy, towards the field of Artificial Superintelligence.

References

[1] T. W. Fong, J. D. Frank, J. M. Badger, I. A. Nesnas, and M. S. Feary, "Autonomous Systems Taxonomy." May 14, 2018. Accessed: Mar. 12, 2023. [Online]. Available: https://ntrs.nasa.gov/citations/20180003082

[2] F. Figueroa, L. Underwood, M. G. Walker, and J. Morris, "NASA Platform for Autonomous Systems (NPAS)," presented at the AIAA SciTech 2019, San Diego, CA, Jan. 2019. Accessed: Mar. 12, 2023. [Online]. Available: https://ntrs.nasa.gov/citations/20180007990

[3] M. Kirshner and E. Pearce, "Developing Optical Sensor Constellation Architectures for Space Domain Awareness through Model-Based Trade Studies," in *Proceedings of the Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*, 2024, p. 42.

[4] C. T. S. I. W. Group and others, "National cislunar science and technology strategy," *National Science and Technology Council, Tech. Rep*, 2022.

[5] S. M. De Gyurky and M. A. Tarbell, *The Autonomous System: A foundational synthesis of the sciences of the mind*. John Wiley & Sons, 2013.

[6] J. F. Figueroa and M. G. Walker, "Integrated system health management (ISHM) and autonomy," in *2018 AIAA Information Systems-AIAA Infotech@ Aerospace*, 2018, p. 1152.

[7] F. Figueroa and P. J. Ramos, "NPAS Physics Pump Models for Autonomous Monitoring," Jan. 01, 2018. Accessed: Dec. 31, 2022. [Online]. Available: https://ntrs.nasa.gov/citations/20180008698

[8] F. Figeuroa, K. A. Woo, and L. Underwood, "Condition-based-maintenance of Electrical Motors for INSIGHT (Intelligent Stennis Gas House Technology) and NPAS (NASA Platform for Autonomous Systems)," Dec. 12, 2019. Accessed: Dec. 31, 2022. [Online]. Available: https://ntrs.nasa.gov/citations/20200000456

[9] F. Figueroa, L. Underwood, and D. Armstrong, "Model-Based Systems Engineering, Real-Time Operations, and Autonomy," in *ASCEND 2020*, in ASCEND. , American Institute of Aeronautics and Astronautics, 2020. doi: 10.2514/6.2020-4091.

[10] F. Figueroa, L. Underwood, B. Hekman, and J. Morris, "Hierarchical Distributed Autonomy: Implementation Platform and Processes," in *2020 IEEE Aerospace Conference*, Mar. 2020, pp. 1–9. doi: 10.1109/AERO47225.2020.9172572.

[11] L. Murdock, F. Figueroa, and L. W. Underwood, "Autonomy Voice Assistant for NPAS (NASA Platform for Autonomous Systems)," in *AIAA SciTech 2021 Forum*, American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2021-0393.

[12] S. Grampone, W. Litwin, and T. Schwarz, "An Autonomous Data Structure for Brute Force Calculations in the Cloud," in *2015 IEEE 7th International Conference on Cloud Computing Technology and Science (CloudCom)*, Nov. 2015, pp. 347–354. doi: 10.1109/CloudCom.2015.17.

[13] G. C. H. E. de Croon, J. J. G. Dupeyroux, S. B. Fuller, and J. A. R. Marshall, "Insect-inspired AI for autonomous robots," *Science Robotics*, vol. 7, no. 67, p. eabl6334, Jun. 2022, doi: 10.1126/scirobotics.abl6334.

[14] "Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices | npj Digital Medicine." Accessed: Mar. 12, 2023. [Online]. Available: https://www.nature.com/articles/s41746-018-0040-6

[15] S. Thombre *et al.*, "Sensors and AI Techniques for Situational Awareness in Autonomous Ships: A Review," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 1, pp. 64–83, Jan. 2022, doi: 10.1109/TITS.2020.3023957.

[16] L. Conway, R. Volz, and M. Walker, "Tele-autonomous systems: Methods and architectures for intermingling autonomous and telerobotic technology," in *1987 IEEE International Conference on Robotics and Automation Proceedings*, Mar. 1987, pp. 1121–1130. doi: 10.1109/ROBOT.1987.1087923.

[17] P. Skobelev, "Towards Autonomous AI Systems for Resource Management: Applications in Industry and Lessons Learned," in *Advances in Practical Applications of Agents, Multi-Agent Systems, and Complexity: The PAAMS Collection*, Y. Demazeau, B. An, J. Bajo, and A. Fernández-Caballero, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2018, pp. 12–25. doi: 10.1007/978-3-319-94580-4_2.

[18] "White House Releases National Cislunar S&T Strategy – Office of Space Commerce." Accessed: Aug. 14, 2024. [Online]. Available: https://www.space.commerce.gov/white-house-releases-national-cislunar-st-strategy/

[19] T. E. Bender, A. S. Gabhart, M. J. Steffens, and D. N. Mavris, "Defining and Parameterizing the Design Space for Cislunar PNT Architectures," in *AIAA SCITECH 2023 Forum*, American Institute of Aeronautics and Astronautics. doi: 10.2514/6.2023-1504.

[20] B. Hagenau, B. Peters, R. Burton, K. Hashemi, and N. Cramer, "Introducing The Lunar Autonomous PNT System (LAPS) Simulator," in *2021 IEEE Aerospace Conference (50100)*, Mar. 2021, pp. 1–11. doi: 10.1109/AERO50100.2021.9438538.

[21] R. Mitch, G. Weaver, J. Bruzzi, W. Millard, B. Summers, and J. Bradfield, "Lighthouses: An Extensible Cislunar Positioning, Navigation, and Timing Architecture," presented at the Proceedings of the 2022 International Technical Meeting of The Institute of Navigation, Jan. 2022, pp. 438–452. doi: 10.33012/2022.18203.

[22] F. Jiancheng and N. Xiaolin, "Celestial Navigation Methods for Space Explorers," *Measurement and Control*, vol. 41, no. 3, pp. 77–80, Apr. 2008, doi: 10.1177/002029400804100302.

[23] S. G. Hesar, J. S. Parker, J. M. Leonard, R. M. McGranaghan, and G. H. Born, "Lunar far side surface navigation using Linked Autonomous Interplanetary Satellite Orbit Navigation (LiAISON)," *Acta*

Astronautica, vol. 117, pp. 116–129, Dec. 2015, doi: 10.1016/j.actaastro.2015.07.027.

[24] M. Kirshner and R. Valerdi, "Integrating Model-Based Systems and Digital Engineering for Crewed Mars Mission Planning," *Journal of Aerospace Information Systems*, vol. 19, no. 10, pp. 668–676, Oct. 2022, doi: 10.2514/1.I010986.

[25] "The Autonomous System: A Foundational Synthesis of the Sciences of the Mind | Wiley." Accessed: Mar. 12, 2023. [Online]. Available: https://www.wiley.com/en-us/The+Autonomous+System:+A+Foundational+Synthesis+of+the+Sciences+of+the+Mind-p-9781118294246

[26] M. Tarbell, "AI and the transcendence of true autonomy," in *Micro- and Nanotechnology Sensors, Systems, and Applications XI*, M. S. Islam and T. George, Eds., Baltimore, United States: SPIE, May 2019, p. 74. doi: 10.1117/12.2517403.

[27] R. Rovetto, *The Orbital Space Environment and Space Situational Awareness Domain Ontology – Toward an International Information System for Space Data*. 2016.

[28] "Human-Computer Interaction for Space Situational Awareness (SSA): Towards the SSA Integrated Sensor Viewer (ISV) | SpringerLink." Accessed: Mar. 12, 2023. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-030-21565-1_34

[29] J. J. Critchley-Marrows, X. Wu, and I. H. Cairns, "An architecture for a visual-based PNT alternative," *Acta Astronautica*, vol. 210, pp. 601–609, 2023.

[30] M. Hartigan, "Simulation and Analysis of Navigation Performance for Cislunar PNT Constellations," 2023.

[31] Z. R. McLaughlin, R. E. Gold, S. G. Catalan, B. A. Jones, and R. Zanetti, "Crater navigation and timing for autonomous lunar orbital operations in small satellites," in *Proceedings of the 44th Annual American Astronautical Society Guidance, Navigation, and Control Conference, 2022*, Springer, 2024, pp. 155–171.

[32] J. Wysack and K. Ferrant, "Contrasting Architectures for Cislunar SDA and STM," in *Proceedings of the Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference*, 2023, p. 207.

[33] S. M. Lehman, N. H. Campbell, S. A. Aytes, M. Kirshner, and A. Arviola, "EnDEVR: An Environment for Data Engineering in VR," in *EnDEVR: An Environment for Data Engineering in VR*, IEEE, 2021.

[34] P. Zimmerman, T. Gilbert, and F. Salvatore, "Digital engineering transformation across the Department of Defense," *The Journal of Defense Modeling and Simulation*, vol. 16, no. 4, pp. 325–338, 2019.

[35] J. M. Badger, P. Strawser, and C. Claunch, "A distributed hierarchical framework for autonomous spacecraft control," in *2019 IEEE Aerospace Conference*, IEEE, 2019, pp. 1–8.

[36] J. M. Badger, C. Claunch, and F. Mathis, "Modular autonomous systems technology framework: A distributed solution for system monitoring and control," 2017.

[37] B. Bowles, M. Hancock, M. Kirshner, and T. Shaji, "Biomimetic Design in Augmented Cognition," in *Augmented Cognition. Theoretical and Technological Approaches*, D. D. Schmorrow and C. M. Fidopiastis, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2020, pp. 226–240. doi: 10.1007/978-3-030-50353-6_17.

[38] J. C. Crusan *et al.*, "Deep space gateway concept: Extending human presence into cislunar space," in *2018 IEEE Aerospace Conference*, IEEE, 2018, pp. 1–10.

[39] J. Crusan *et al.*, "NASA's gateway: An update on progress and plans for extending human presence to cislunar space," in *2019 IEEE Aerospace Conference*, IEEE, 2019, pp. 1–19.

[40] K. J. Runnels and D. Gebre-Egziabher, "Gravitational Acceleration Maps for Initialization and Ambiguity Resolution of Pulsar-Based PNT in Space," in *Proceedings of the 35th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2022)*, 2022, pp. 1917–1932.

[41] F. Figueroa, L. Underwood, J. Morris, M. Walker, and R. Brown, "Risk-Reduction Autonomy Implementation to Enable NASA Artemis Missions," in *2022 IEEE Aerospace Conference (AERO)*, 2022, pp. 1–13. doi: 10.1109/AERO53065.2022.9843815.

[42] F. Chollet, "On the measure of intelligence," *arXiv preprint arXiv:1911.01547*, 2019.

[43] M. Kirshner, "Achieving Holistic Interoperability with Model-Based Systems Engineering," PhD Thesis, The University of Arizona, 2023.

[44] R. Chatila and J. C. Havens, "The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems," in *Robotics and Well-Being*, vol. 95, M. I. Aldinhas Ferreira, J. Silva Sequeira, G. Singh Virk, M. O. Tokhi, and E. E. Kadar, Eds., in Intelligent Systems, Control and Automation: Science and Engineering, vol. 95. , Cham: Springer International Publishing, 2019, pp. 11–16. doi: 10.1007/978-3-030-12524-0_2.

[45] "Defense Innovation Board Recommends AI Ethical Guidelines," U.S. Department of Defense. Accessed: Mar. 12, 2023. [Online]. Available: https://www.defense.gov/News/News-Stories/Article/Article/2006646/defense-innovation-board-recommends-ai-ethical-guidelines/se-innovation-board-recommends-ai-ethical-guidelines/

[46] *Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts*. 2021. Accessed: Mar. 12, 2023. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52021PC0206

[47] I. Kant, "Critique of pure reason. 1781," *Modern Classical Philosophers, Cambridge, MA: Houghton Mifflin*, pp. 370–456, 1908.

[48] "NASA TechPort - Project." Accessed: Jan. 12, 2025. [Online]. Available: https://techport.nasa.gov/project

[49] R. Ambrose *et al.*, "NASA technology roadmaps: TA 4: Robotics and autonomous systems," *NASA, Washington DC*, vol. 18, 2015.

[50] M. Kirshner, "Model-Based Systems Engineering Cybersecurity for Space Systems," *Aerospace*, vol. 10, no. 2, Art. no. 2, Feb. 2023, doi: 10.3390/aerospace10020116.

[51] A. Domotor, M. Borowitz, and R. Palmer, "Cislunar Positioning, Navigation, and Timing: International Relations and Policy Implications," *New Space*, vol. 11, no. 4, pp. 251–261, Dec. 2023, doi: 10.1089/space.2023.0034.

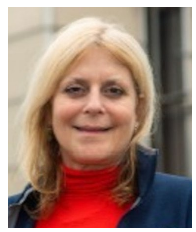
[52] B. Horowitz, P. Beling, K. Skadron, R. D. Williams, and W. Melvin, "Security engineering project," *Technical Report SERC-2015-TR-036-4*, 2015.

[53] J. Bayuk, B. Horowitz, and R. Jones, "Security via related disciplines," *Procedia Computer Science*, vol. 8, pp. 338–344, 2012.

[54] B. Horowitz *et al.*, "Cyber security requirements methodology," *Systems Engineering Research Center, Technical Report SERC-2018-TR-110*, 2018.

BIOGRAPHY

Mitchell Kirshner *is a postdoctoral research associate at Steward Observatory at the University of Arizona. He finished his PhD in Systems & Industrial Engineering at the University of Arizona in May 2023. His specialty is in Model-Based Systems Engineering (MBSE), a field which uses semantic interoperability and visual programming languages to form cross-cutting diagrams as an alternative to document-based engineering of complex system-of-systems. In his research, Dr. Kirshner utilizes MBSE to advance space telescope constellation optimization through requirements-based trade studies. Prior to his enrollment at UA, he worked at submarine manufacturer General Dynamics Electric Boat and attended Northwestern University for his BS and MS. He is originally from Long Island, New York.*



Lauren Underwood *received her BS, MS, and PhD Degrees in Biology (developmental neurobiology with focus in optics) from Tulane University, LA (1985, 1987 & 1991). Following graduation, she was awarded a National Institute of Health Training Grant Fellowship in Vision Research for postdoctoral studies (1991–1993). Dr. Underwood has been at NASA Stennis Space Center (SSC) since 2001, and has worked directly for the Office of the Chief Technologist (OCT), as well as supported activities associated with Science Mission Directorate, Science Technology Mission Directorate (STMD), DEVELOP and most recently works in the Test Technology Branch in the Engineering and Test Directorate at SSC. Currently, she serves as an SSC Project Manager, for NASA's Advance Exploration Systems (AES) autonomy software activities, as well as serves as the Autonomous Systems Lab project manager for other autonomy software infusion activities at SSC.*

Fernando Figueroa *is the lead principal investigator in projects involving multiple NASA centers, universities, laboratories, and industry. His professional experience encompasses a combination of management, leadership, and strong technical credibility (NASA-wide and outside NASA) needed to provide focused and purposeful direction for the successful execution of projects and programs. He has developed research programs that merited funding from federal governments (United States, and Canada), state governments, and private industry. He is a Subject Matter Expert (SME) on autonomous systems and operations, and a member of NASA working groups and study groups addressing autonomy requirements, concepts of operations, and technology to support Artemis, planetary exploration, and ground test/launch systems.*