Leto™ - A Spaceflight Intelligence System

Jake Rohrig[1], Taylor Kirk[2], Miguel Torres[3]
Collins Aerospace, Windsor Locks, CT, 06096-1010 USA

Hunter Sandidge[4], Jackson Wehr[5], Jacob Werschey[6]
Collins Aerospace, Rockford, IL, 61108-7929 USA

As the commercial space industry accelerates and humanity sets its sights on deep-space destinations, the human capital required to support these bold endeavors grows dramatically. To alleviate the strain on the industry workforce, reduce cost, and optimize operations, Collins Aerospace (an RTX business) is developing a spaceflight intelligence system, Leto™, that integrates with existing and future systems. Leto™ is a full-stack system that uses physics-based models and AI-powered algorithms to continuously monitor and inform ground support personnel and onboard crew about the performance of their vehicle's Environmental Control and Life Support System (ECLSS). Leto™, trained and verified on decades of spaceflight telemetry, surveys a vehicle's ECLSS and provides performance metrics, prognostics, and anomaly detection functions that alert users to system degradation and advise when upcoming maintenance events should occur. Intelligent ECLSS monitoring with Leto™ allows greater insight into system performance while reducing the labor required to do so, letting critical engineering staff focus on value-added activities. This paper introduces Leto™, discusses its capabilities through actual use cases, and concludes with next steps.

Nomenclature

AI	=	Artificial intelligence
ECLSS	=	Environmental control and life support systems
EMU	=	Extravehicular mobility unit
ETL	=	Extract-Transform-Load
EVA	=	Extravehicular activity (e.g. spacewalk)
HCI	=	Human Computer Interface
ISS	=	International Space Station
KPI	=	Key performance indicator
LLM	=	Large Language Model
MFB	=	Multifiltration Bed
ML	=	Machine learning
NASA	=	National Aeronautics and Space Administration
PFD	=	Process flow diagram
PIML	=	Physics informed machine learning
R&R	=	Removal and replacement
RUL	=	Remaining useful life
SME	=	Subject matter expert
vSME	=	Virtual subject matter expert
WPA	=	Water Processor Assembly
PFD	=	Process flow diagram
PIML	=	Physics informed machine learning
R&R	=	Removal and replacement
RUL	=	Remaining useful life
SME	=	Subject matter expert
vSME	=	Virtual subject matter expert
WPA	=	Water Processor Assembly
xEVAS	=	Exploration Extravehicular Activity Services

[1] Space Technology Discipline Chief, Mission Systems Advanced Technology & Capabilities, 1 Hamilton Road, Windsor Locks, CT 06096-1010.
[2] Mgr. User Experience Design, Connected Aviation Solutions, 1 Hamilton Road, Windsor Locks, CT 06096-1010.
[3] Sr. Principal Eng, Data Analytics & Methods, Advanced Research & Technology, 1 Hamilton Road, Windsor Locks, CT 06096-1010.
[4] Mgr. Data Science & Artificial Intelligence Engineering, Mission Systems, 4747 Harrison Avenue, MS 290-6, Rockford, IL 61108.
[5] Senior Data Scientist, Mission Systems, 4747 Harrison Avenue, MS 290-6, Rockford, IL 61108.
[6] Sr. Data Engineer, Mission Systems, 4747 Harrison Avenue, Rockford, IL 61108.

I. Introduction

Over the past decade, the space economy has expanded by an astounding $260 billion, with commercial revenue constituting 78% of that market. This growth has been fueled by technological advancements, private investments, and a surge in interest from both established players and innovative startups.[1] As companies and agencies race to launch low-earth orbit stations, explore lunar resources, and pioneer deep-space travel to new destinations, the landscape of space-related activities and quantity of diverse missions is rapidly growing. This rapid expansion of concurrent missions will strain the spaceflight industry workforce and, as a result, challenge existing methods and frameworks of providing mission oversight. This phenomenon is depicted in Figure 1.

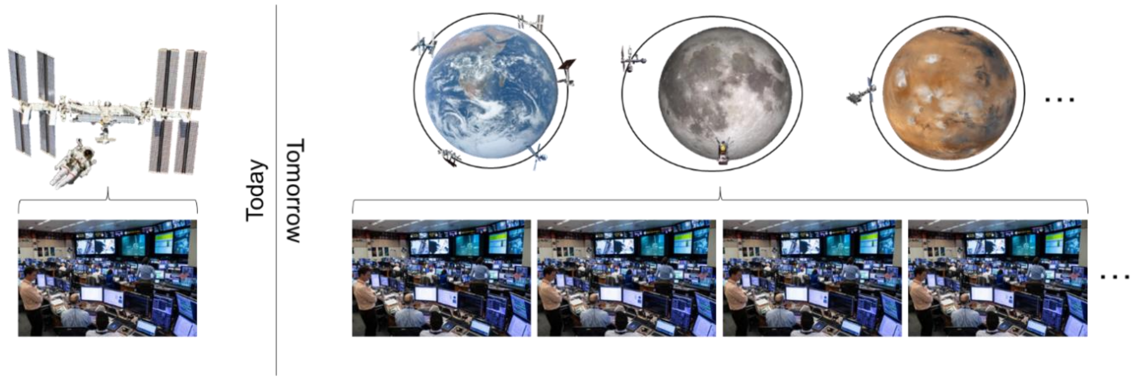


Figure 1. A graphical depiction of oversight methodologies today and the scaling required to sustain continued use of those methods in an expanding space economy.

In human spaceflight operations today, much of the supporting analytics are performed manually and include extensive human-in-the-loop processes. These methods, while proven and robust, add sustaining mission costs and require low-supply, highly-specialized human capital. Ultimately, the current methodology for analytics and mission support will not be able to fully address the future growth and challenges of the industry; the human spaceflight industry demands novel analytical tools that can handle large-scale data, optimize resource allocation, and assess risk in an ever-changing environment.

In this paper, Collins Aerospace (an RTX business) will expand in greater detail on the challenges experienced with current oversight and analytics approaches to monitoring vehicle Environmental Control and Life Support Systems (ECLSS), present a vision for an ideal state of ECLSS monitoring and mission oversight, and introduce Leto™, Collins' Spaceflight Intelligence System. In the latter half of this paper Leto™ will be described in greater detail and through example, show how it can be used to overcome the current expected mission support challenges induced by a rapidly expanding spaceflight industry.

II. Current Domain Challenges

No goal, especially one as ambitious as human space exploration, comes without challenges. When it comes to using data to support human spaceflight, the challenges are great. The following subsections describe some of the most common and intrusive challenges faced in this domain.

A. Data Challenges

First, the sheer quantity of telemetry poses a set of challenges. Each system, subsystem, and sensor have unique functional bounds, states of operation, and telemetry delivery systems which create an immense volume of unstandardized and unstructured data across many disparate sources. Consequently, to properly monitor the health and performance of space hardware, an incredible amount of time is spent aggregating, cleaning, standardizing, and storing data before any meaningful analysis can even begin. Additionally, the varying sample rates of data collection (e.g., one sample per second vs one sample per minute) make cross-comparison of data sets challenging. In some scenarios, the data collection rate is so high that even the best analytic and visualization software struggle due to the volume and noise that tends to accompany high frequency data. Yet, in other scenarios, the data collection rate is so low that important dynamic behaviors that happen in small windows of time go unrecorded and valuable performance data is lost.

B. Human Capital Challenges

The problems stated above are exacerbated by a limited supply of human capital both capable and with the domain expertise to effectively perform data analysis and monitoring. This data-versus-human-capital-supply disparity leads to a host of problems including, but not limited to:

a) a lack of time to improve or advance existing monitoring methodologies
b) monitoring only a subset of all components that should be monitored
c) human error in manual data entry
d) human error in calculations
e) untrustworthy data sets
f) suboptimal use of specialists' time (e.g., performing low value-added work required before beginning meaningful diagnostics)
g) not being current in the analysis of the most recent data (inability to monitor data in real-time)
h) operating reactively (i.e., only analyzing things after they have gone wrong)
i) a lack of time to train additional human capital, therefore risking loss of knowledge through attrition
j) a lack of process documentation, standard operating procedures, and other explanatory process information

Additionally, the background of those currently monitoring spaceflight hardware is primarily in the engineering discipline and not typically in the data science domain; this discrepancy in skills leads to missed opportunities to leverage newer or more advanced health monitoring techniques, such as artificial intelligence (AI) and machine learning (ML) based models, as well as data analytics automation through software development. Relatedly, because each system producing telemetry has a different function (e.g., water processing, oxygen generation, etc.), a specialized engineer is often dedicated to understanding and monitoring each system. As experience is gained, this engineer becomes a subject matter expert (SME) and the specialization creates analytical silos, reducing the overall ability to analyze and understand system interactions of an ECLSS. The specialization also often creates organizational single points of failure and dependencies on an individual's short and long-term availability.

C. Analytics Challenges

Outside of the aforementioned problems, additional issues result from the underlying nature of the telemetry-generating hardware and the dynamic nature of operating that hardware. For example, most flight hardware has never seen or experienced an in-flight failure. Without failure data, it is infeasible to build any non-analytical (i.e., AI or ML) predictive models given that the events to be predicted have never happened, and therefore do not have signatures in the data to be used for AI/ML training. Even consumables (e.g., filters) that have finite life and do have on-orbit replacement histories for AI to learn from, are not often run to failure and are changed while still healthy. This is typically done to prevent a failure mode from causing larger system problems, limited availability of launch schedules, or simply for convenience and efficient use of time onboard the spacecraft. This practice dilutes the purity of the data such that only removal and replacement (R&R) events, which may or may not be failure-driven events, are observable in the data. Relatedly, Mission Control frequently changes operating conditions on various systems and subsystems, such as changing allowed operating bounds. This modification in operating conditions over time creates significant and frequent structural breaks in the data. These structural breaks also add to the impurity of the historic data, thereby making any anomaly detection or prognostics more challenging to develop from these data sets. This challenge is exemplified in Section V, where a prognostic is presented which predicts the first ionic breakthrough of the International Space Station (ISS) Water Processor Assembly (WPA) Multifiltration Bed (MFB).

III. Ideal State Solution

Given the problems and challenges presented in the previous sections, one can imagine how an intelligent system might be used to overcome them and what the solution might look like for the purposes of monitoring spacecraft life support systems for human spaceflight. This ideal solution's end-to-end description of capabilities can be broken into four major categories. We'll describe these categories in sequential order, following the path of data, from the telemetry generating hardware to the delivery of valuable information to the user. These categories of capabilities are:

a) Data collection, curation, and supplementation
b) Analytics
c) Human computer interface (HCI)
d) Deployment

D. Data Capabilities

Starting with raw telemetry ingestion, an ideal solution would have a standardized, relationally linked database containing all hardware telemetry history and metadata surrounding the telemetry. An intelligent database design that is product agnostic would be essential to allow for the rapid addition of new hardware without any need to change upstream code to perform analytics or serve a front-end application. The centralization and aggregation of historic telemetry is essential to create a single source of truth as well as a deep repository for the sake of building AI and ML models. This database would also need the ability to ingest data in real-time to allow for streaming analytics and health monitoring.

In addition to a product-agnostic standardized database design, the aggregation of all historic telemetry, and the ability to ingest new data in real-time, a secondary set of synthetic data would also be required. Given that most hardware has never seen failure modes or been in service long enough to see the end-of-life performance degradation, it is infeasible to train AI to identify or predict failure states or product lifespans without the exemplary data. Thus, there is a need for a digital twin that could be used to intentionally cause, observe, and collect data on failure modes. This synthetic data repository, in conjunction with and in similar design as the real telemetry database, would be essential for building proper AI training data sets to overcome the current *rare event* problem.

Unlike the current process for many NASA contractors, where raw data is manually downloaded, extracted, transformed, and loaded into spreadsheets for data collection and analysis, the ideal state would have a direct connection to the raw data source such that a seamless, automated pipeline would exist between either the on-ground telemetry receiving server or the telemetry-generating hardware and the proposed database. This would allow for on-the-fly standardization, cleansing, and transformation such that streaming analytics could then be performed in real time. Additionally, this would prevent human error in the extract-transform-load (ETL) process.

E. Analytics Capabilities

After data ingestion, a layer of standardized and automated reporting would need to exist on top of the proposed database. Creating a suite of standardized health monitoring methods and metrics would allow for common treatment, assessment, and interpretation of product health, regardless of hardware or platform. Additionally, this standardization would decrease the burden of learning for any engineer tasked with monitoring the hardware as it would be in common form. This common suite of health metrics would need to include such things as anomaly detection, breaches of expected telemetry bounds, and prognostics for remaining useful life (RUL) forecasts. Forward-looking metrics would need to be included within this suite to allow SMEs to work proactively to prevent problems, instead of reactively diagnosing them.

This layer of common analytics would need to be automated such that the analytics are generated and provided immediately upon data ingestion. Automated system alerts would also be required for these common health measuring metrics to help the engineering staff focus their limited time on core issues. This automated system would need to have triggers that would automatically induce all system ML models to retrain based on new information. For example, if a piece of hardware had failed and was replaced, then that new information, once passed into the intelligent monitoring system, would automatically retrain and update a RUL ML model such that the model would have one additional event in its training history to understand how to predict product lifespan. Complementary to automated retraining from triggered events, the system would also have automated ML model drift monitoring. The system would need to constantly self-validate that its ML models were current and that its performance matched the expected performance from its training period. If a model was found to have drifted, it would also need to be able to retrain itself.

Finally, the analytics layer would need to be modular, allowing for inclusion of idiosyncratic analytics specific to the hardware under supervision. Each piece of hardware monitored in spaceflight has key performance indicators (KPIs) unique to its performance and health. The ability to add these specialized metrics in an unobtrusive manner to both backend and frontend of the intelligent system would be paramount for rapid hardware analytics deployment and critical for meaningful performance analysis.

F.Human Computer Interface Capabilities

Once system monitoring information has been derived from the data with the analytics layer, the ideal solution needs to have a way to interact with and add value to the user. This will require changes to the tools and processes people use today to monitor spaceflight equipment. As previously mentioned, the ideal solution would remove the burden of low-value-add work (data downloading, cleansing, aggregation, chart generation, metric calculation, etc.) from the limited supply of engineering subject matter experts (SMEs). The SMEs' time should be optimized such that they only work on relevant issues and perform deeper analyses. Automation of these tasks in conjunction with the predictive and pointed alert systems would largely contribute to optimizing the SMEs' time.

Additionally, the solution would need to codify SMEs' domain knowledge to create an embedded and reviewable knowledge base. This knowledge base would serve as a structured form of documentation to prevent knowledge loss under scenarios of attrition and could be used to both train new employees and cross-train existing employees. The ideal state would aggregate, in real time, all corpus related to product history and health analysis such that an AI virtual SME could be built and updated. This virtual SME (vSME) would then expand capacity of human SMEs by being able to train, answer questions, etc. in place of the human SME. The vSME would be able to serve as a point of first contact before a human SME would need to be involved, and it would retain the entire history of all SMEs interacting with the system, making it in a sense a super SME composed of domain knowledge from many human SMEs both through time and across hardware functions. The vSME could also frequently ingest new information relevant to the monitored hardware to add to the knowledge base outside of the human SMEs purview.

A final human computer interface capability in the ideal solution would allow for a human-in-the-loop feedback mechanism, such that new information could be added to the system to kick off the aforementioned automated events. For example, if a consumable on orbit was changed, the human SME could log this event, which would then kick off automated ML retraining. Additionally, this mechanism could enable crowdsourcing of information and event labelling for the sake of building AI models.

G. Deployment Mechanism Capabilities

The last layer of the ideal solution is the delivery mechanism. An ideal content delivery system would be able to show all insights, metrics, etc. in a singular, unified platform such that all monitored hardware could be reviewed in one destination, with the ability for users to query and dig deeper into any insight they may wish to review. This centralization of all cointegrated hardware would allow for a synergistic and cross-function understanding of hardware health. The user interface would need to be able to be served locally (e.g., desktop), on mobile devices, and at the point of use (i.e. on the edge) to provide better alerts, faster reactions, and improved performance for both crews and mission support personnel. A final capability that the user interface should have is the ability to be modular, such that new products and analytics could be added rapidly without the need for costly user experience overhauls.

IV. Introducing Leto™, A Spaceflight Intelligence System

To overcome the issues in the space analytics domain, our team created Leto™, a spaceflight intelligence system, guided by the vision posed in Section III. "Leto" was chosen as the name for our intelligent system because in Greek mythology, Leto is the daughter of Coeus (the Titan of intelligence and wisdom) and Phoebe (embodying lunar radiance, wisdom, and prophetic insight), personifying the joining of rational intellect and prophetic wisdom. Leto is also the mother of Apollo and Artemis, a subtle nod to key human exploration programs of the past and future.

Leto™ is a full-stack, standardized, and scalable AI and analytics system. It is capable of both ex-post and streaming analytics designed for rapid integration of new products with the overarching goal to provide common capabilities to all Collins' spaceflight hardware (e.g., ECLSS, power management and distribution systems, etc.) across all platforms (e.g., International Space Station, Orion, Space Launch System, spacesuits, etc.). Leto™ was made to:

a) automate data aggregation, preprocessing, and analysis
b) standardize analytic investigation, presentation, and interpretation
c) automatically detect and alert engineers of anomalous behavior
d) predict remaining useful life of hardware
e) expand engineering SME impact by eliminating low-value-add activities and directing them to key problems
f) serve as a medium for SME knowledge capture and on-demand knowledge retrieval as a vSME

The system, in its current form, is deployed for ground-based mission support through a web application, depicted in Figure 2. The capabilities of the web application can be divided into categories, *core* and *idiosyncratic*, respectively.

A. Core Capabilities

The core capabilities are offered across all product lines embedded in the Leto™ spaceflight intelligence system. In its current form, these core capabilities include:

- *Telemetry Processing* - the ability to view, filter, and download raw telemetry
- *Process Flow Diagrams (PFD)* – to allow users to explore the process flow diagram of the system under observation to better understand the telemetry that is being captured and monitored
- *Analytics* - view, filter, and download statistics across all observable parameters and includes:
    - Sensor Drift Detection - a novel mathematical method that scores an individual sensor's performance between 0 and 1, providing an interpretable value for nominal/anomalous behavior
    - Breach Detection – minimum and maximum breach detection set across all sensors
- *Prognostics* – estimates of the remaining useful life of hardware through applied machine learning
- *Anomaly Detection* - detection of performance anomalies through applied machine learning models that flag abnormal behavior of hardware and subsystems

B. Idiosyncratic Capabilities

Given that each spaceflight product has unique build and function, it is imperative that the Leto™ spaceflight intelligence system allows for inclusion of idiosyncratic KPIs for each product being monitored. For example, if viewing the International Space Station's Water Processing Assembly, one would want to see metrics related to water production, available quantity, etc., whereas if viewing performance of an extravehicular mobility unit (EMU) during a spacewalk, one would want to see parameters like remaining oxygen, remaining battery, suit temperature and pressure, etc. Recognizing this need, the Leto™ spaceflight intelligence system has been designed to be modular, allowing for each product line to have its own collection of specialized KPIs as well as the ability to rapidly add and update them for visualization at the user interface.

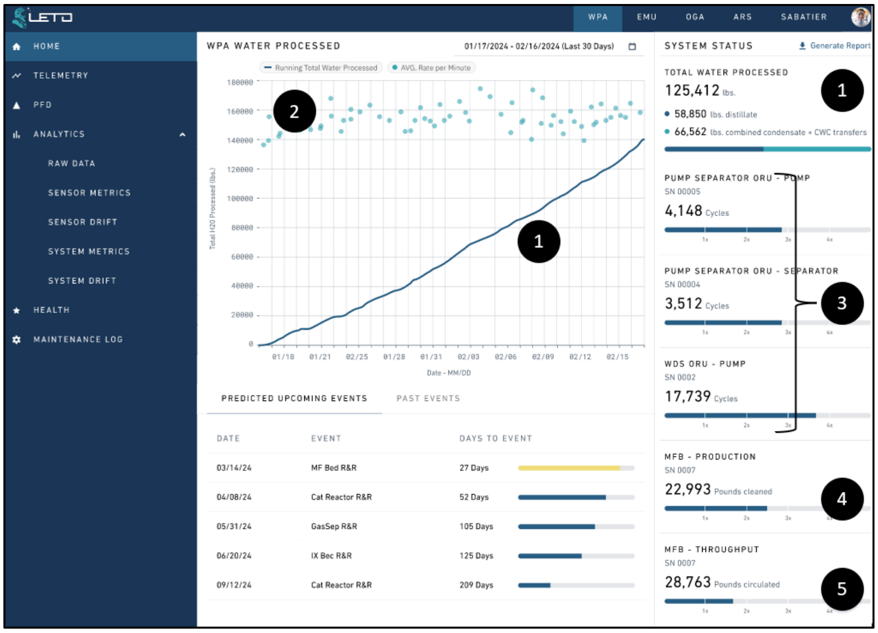
V. Leto™ Through Example

To better illustrate some of the capabilities and value offered by Leto™, examples of how the system can be used to monitor the water processor assembly (WPA) onboard the ISS is presented below.

A. WPA Idiosyncratic Capabilities (KPIs) Example

Beginning with idiosyncratic capabilities, our development team worked with the engineers responsible for supporting and monitoring the ISS WPA to capture the KPIs that are needed to understand the performance and health of the WPA. After a review of the KPIs and methodologies, our team concatenated a lifetime data set from disparate historic data files, developed a data ingestion pipeline for future data collection, and created all relevant WPA KPI monitoring algorithms (in addition to entirely new KPIs desired by WPA SMEs). Some of the KPIs that are tracked and analyzed via the Leto™ web application are shown in Figure 2. By automating the data collection and monitoring of KPIs, our team has been able to streamline the process of KPI information generation, allowing SMEs to focus on interpretation of the information rather than generation and documentation.

While SMEs have found general value in offloading the work required to prepare and monitor the WPA KPIs, there have been several unforeseen benefits and uses of Leto™ capabilities that are worth discussing. One observed benefit is the ability to provide SMEs with lifetime data sets for the WPA; by centralizing and concatenating the raw data from the WPA, engineers can now perform large-scale studies in a manner previously impracticable. Prior to Leto™, each day's worth of WPA data was stored and archived as an individual file, making long-term trending analysis infeasible. If an engineer wanted to study the performance deterioration of a pump over its entire multiyear life, they would have to open and aggregate hundreds of Excel files. Further complicating matters, Excel can only handle so many days of data before it reaches its maximum number of records. As such, long-term studies were impossible. However, because of core Leto™ capabilities, the engineering staff is now able to see the entire life of any WPA component, helping to substantiate design decisions and advancements in the next-generation WPA.



KPI	
1	Total Water Production
2	Average Daily Production Rate
3	Total Rotating Equipment Spin Up Cycles
4	Total Water Produced – with Multifiltration Bed
5	Total Water Throughput – through Multifiltration Bed

Figure 2. A labeled screen capture of the Leto™ web application indicating ISS WPA KPIs.

B. WPA Prognostics Example – Multifiltration Bed Remaining Useful Life Estimate

The ISS WPA produces potable water from onboard waste water, such as humidity condensate and urine distillate. The primary treatment process is achieved in the WPA Multifiltration Bed (MFB), which includes an adsorbent media and ion exchange resin for the removal of dissolved organic and inorganic contaminants. The MFB removes dissolved organic and inorganic contaminants from the waste water to be processed back into potable water, but its performance degrades as it becomes saturated. When the ion exchange resin becomes saturated with ions, such as calcium, magnesium, or other dissolved salts, it can no longer exchange ions with the water passing through. This saturation leads to what is known as *ionic breakthrough*, where the contaminants start appearing in the treated water downstream from the MFB, which poses a risk to the WPA Catalytic Reactor. To detect ionic breakthrough, a conductivity sensor is placed downstream of the MFB. Increased conductivity can signify that the ion exchange resin in the MFB is no longer effective, and the bed is reaching its capacity. Consequently, patterns in conductivity signifying ionic breakthroughs are indicative of when a MFB requires replacement. [2-3]

This critical component of the WPA is one of the main contributors to the total resupply mass required to sustain operation of the WPA on orbit. Weighing roughly 110 pounds, premature removal and replacement of the MFB is costly. Recognizing the significant value that accurate MFB ionic breakthrough predictions could provide, our team developed a machine learning model to predict when the first ionic breakthrough will occur.

It is worth noting that historically (before July 2019 when the WPA began running with a single MFB [new configuration] and a new sorbent media [Ambersorb] mid-2020)[3] the MFBs (quantity 2) were rotated and replaced due to *organic* breakthrough, not *ionic* breakthrough as intended. However, with the new-sorbent, single-bed configuration in place, along with the installation of siloxane scrubbing air filters in the ISS, the MFB has been allowed to experience multiple ionic breakthrough events before the filter needs to be changed, specifically requiring the change after the third ionic breakthrough.[3] Subsequent breakthrough events present similarly to the first ionic breakthrough, with each successive breakthrough significantly increasing the conductivity of the water downstream of the filter. At present, the Leto™ team has only developed an RUL model indicating the first ionic breakthrough and is working on developing models for both secondary and tertiary breakthroughs. As mentioned in Section II, inconsistent operation and configuration changes dilutes the purity of the data sets available to develop ML RUL models; because of the historic operating conditions of the MFB, there are not many secondary and tertiary breakthrough events to study, which means there is not yet enough example data for a ML model to be built. However, as described in Section VI, Next Steps for Leto™, our team is working on digital twins to synthesize data for ML to fill this technical gap.

As an example of Leto™ prognostic performance, consider that the MFB was replaced on April 29th, 2022, after experiencing a third ionic breakthrough, at which point Leto™ began predicting the days until the newly installed

MFB would experience its first ionic breakthrough. As shown in Table 1 and Figure 3, Leto™ was able to predict with reasonable accuracy the breakthrough event that took place on November 2nd, 2022.

Date of Prediction	RUL Prediction (# of days)	Est. First Ionic Breakthrough Date	Error (# of Days)
5/1/2022	215	12/2/2022	30
6/1/2022	188	12/6/2022	34
7/1/2022	154	12/2/2022	30
8/1/2022	102	11/11/2022	9
9/1/2022	72	11/12/2022	10
10/1/2022	45	11/15/2022	13
11/1/2022	2	11/3/2022	1

Table 1. MFB RUL Model Performance

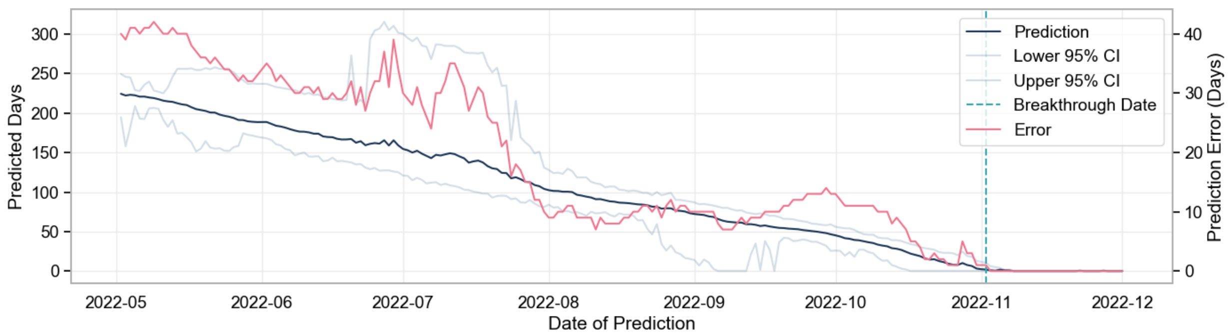


Figure 3. Plot showing the WPA MFB RUL prediction (in number of days remaining until first ionic breakthrough) at each day in 2022.

C. WPA Sensor Drift Detection Example – Multifiltration Bed Ionic Breakthrough

In conjunction with the predicted ionic breakthrough using the RUL model, Leto™ was able to detect the ionic breakthrough using a novel *sensor drift* model. Many of the methods used to detect out-of-family observations rely heavily upon Gaussian assumptions and mean/standard deviation relationships. These approaches tend to suffer from two distinct problems. First, many observed phenomena are not Gaussian, and therefore an attempt to normalize the data must first be performed (e.g., Log, Box-Cox, Yeo-Johnson, etc.), which may or may not actually create normality. Second, assuming the underlying data is perfectly normal, there can be a difference between that which is *statistically* significantly different and that which is *materially* significantly different. Depending on sample size and standard deviation, a tiny change from the mean value may be found to be statistically significant. However, this change may not actually be a difference in magnitude that is of concern for the observed phenomena. Consequently, we set out to devise a new univariate approach that does not suffer from these problems and to score how far an observation is for a sensor based on expected nominal performance.

This model is a custom, proprietary mathematical method that scores an individual sensor's performance between 0 and 1, with 0 meaning the sensor is performing completely nominally and 1 meaning the sensor is behaving abnormally. This sensor drift metric in combination with the auto-alert functionality integrated as a Leto™ core capability has proven useful, autonomously detecting and alerting to unexpected behavioral changes. As shown in Figure 4, this detection method confirmed the ionic breakthrough of the MFB and can be seen in the sensor drift score, attributable to a change in downstream conductivity sensor measurement (also displayed). Progressing through the plot chronologically, the sensor drift score experiences a dramatic increase towards the end of October 2022 and confidently flags the anomalous behavior (ionic breakthrough) on November 2nd, 2022.

In this example it is noteworthy that the drift score is highly correlated with conductivity, as can be visually seen in Figure 4. One will note that conductivity and the drift score spiked June through July. Our team is still in the process of establishing rules for state classification based on drift score, but at present, drift scores less than .6 are considered nominal, values between .6 and .8 are considered possible drift events worthy of investigation, and values exceeding .8 are considered to have drifted. While single point observations are important, especially in real-time, experimentation with the novel drift score method have thus far indicated that trended values over time may be stronger indicators of true drift than any single observation.

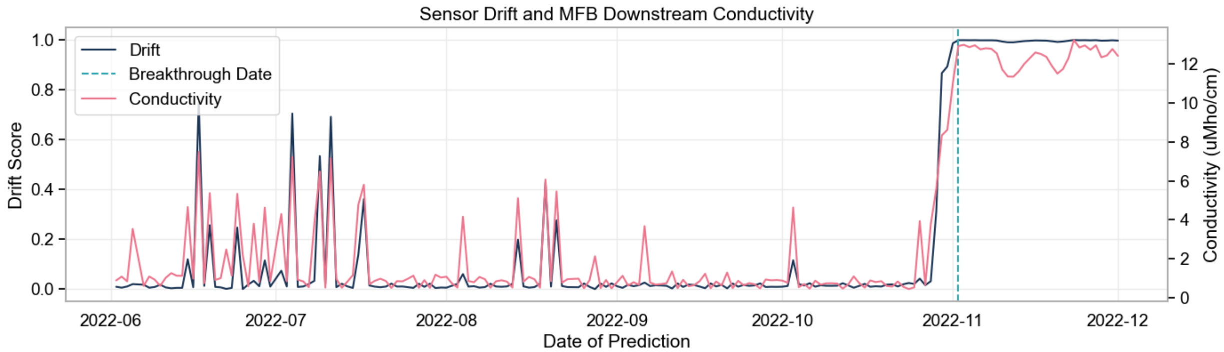


Figure 4. Sensor Drift Anomaly Detection plot of WPA conductivity sensor and conductivity values flagging increasingly anomalous behavior at the end of October 2022.

D. WPA Machine Learning Anomaly Detection Example – Multifiltration Bed Ionic Breakthrough

In addition to the statistical methods described previously, Leto™ is also equipped with *system-level machine learning anomaly detection*. While our previously described anomaly detection methods focus on tracking individual sensors, this capability monitors the system as a whole. For Leto™, our team invented an entirely novel unsupervised machine learning multi-model for the purposes of *explainable* anomaly detection.

Our ML multi-model is tuned to generate a probability that the performance of the system on a given day is anomalous. An output example of this anomaly detection method is depicted in Figure 5, where you can see the relationship between (an algorithm that predicts) current ionic state (i.e., how many ionic breakthroughs have cumulatively occurred until the day of assessment, represented in the legend as integers) and the anomaly probability estimate. The model in Figure 5 was tuned using periods of no ionic breakthrough as nominal, and therefore any period of ionic breakthrough should appear anomalous. If we review a similar period of time as the RUL and sensor drift scores above, you can see the anomaly probabilities (moreover, the variance of the probabilities) change when the filter was replaced on April 29th, 2022, and then again when the newly-installed MFB experienced its first ionic breakthrough on November 2nd, 2022.

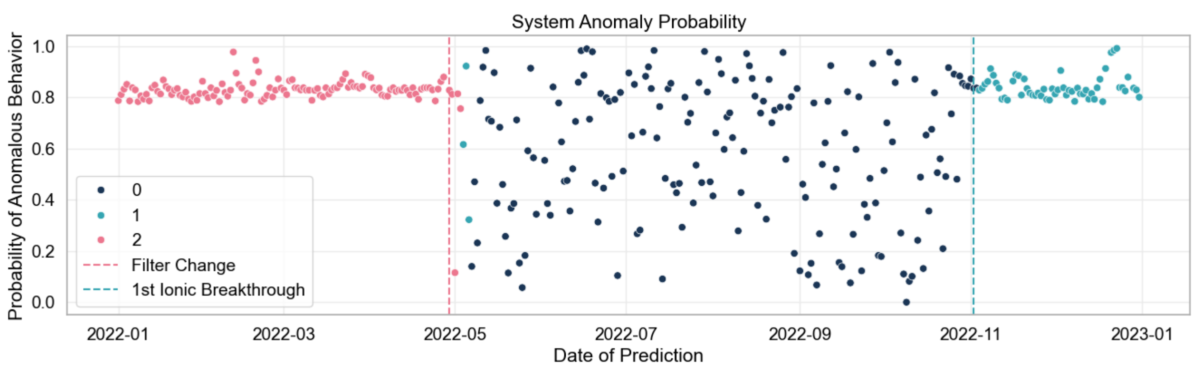


Figure 5. Probability of anomalous system behavior predicted by the explainable system-level anomaly detection ML model.

One of the more significant, novel advancements of our anomaly detection ML method is its *explainability*. Most mainstream anomaly detection models provide a classification, probability, or other score stating whether the observation in question is anomalous. However, these models do not provide information as to why the resultant output is considered anomalous. To help the users of Leto™ focus their investigation when an anomaly is found, a novel methodology for anomaly detection was created. In addition to the forecasted probability score seen in Figure 5, our model also calculates multiple different scores to help the users understand what parts in the system are causing the provided anomaly score.

One such output is the system deviation from a learned nominal state, as presented in Figure 6A. During training, the model learns what the nominal state of the system is and learns a *nominal polygon* with each vertex representing a sensor in the system. During prediction, the input telemetry forms an *observed polygon*. The more dissimilar the observed polygon is from the nominal polygon, the more anomalous the observation. The difference between each observed vertex and its nominal vertex indicates the extent to which the sensor deviates from nominal, and a similarity score between the two polygons is also calculated. This two-dimensional learned representation allows for the entire system to be visually assessed by the engineers to understand which sensors and relationships between sensors drove the predicted anomaly probability.

Additionally, another metric, displayed in Figure 6B, is calculated during the prediction phase to help explain the anomaly probability score. The model learns the nominal *relationships* between all sensor interactions during training. These relationships are then assessed during observation. The model can identify which individual sensors caused deviation from the expected relationships, and then outputs a relational contribution score for each sensor. This score can be interpreted as the percentage that each sensor contributed to the predicted anomaly probability score.

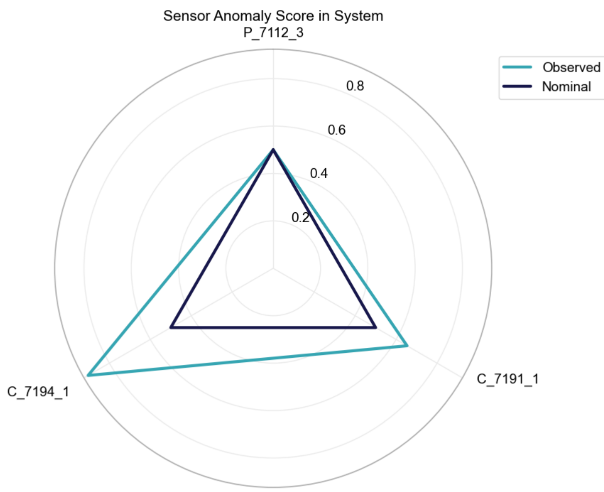


Figure 6A. Polygonal explanatory representation for November 22nd, 2022

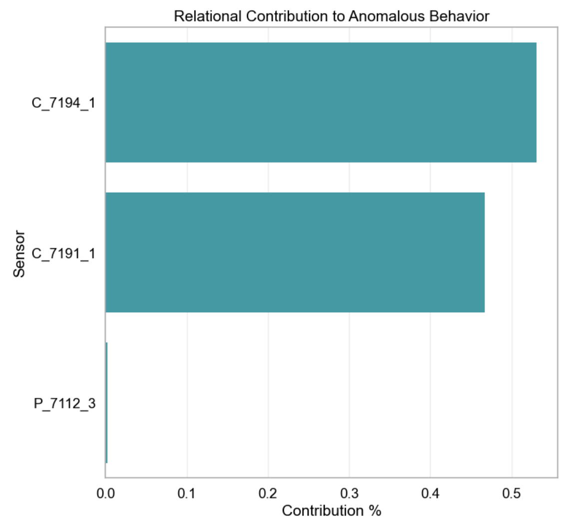


Figure 6B. Relational anomaly contribution by sensor for November 22nd, 2022

VI. Next Steps for Leto™

The Leto™ mission is ambitious in scale and is anticipated to be a multiyear endeavor. While Leto™ in its current form is designed to support ground-based mission control for the ISS's ECLSS, Leto™ will expand its capability to support Collins-provided ECLSS for the commercial space market. In addition, we expect to use Leto™ for spacesuit monitoring of both the existing EMU and Collins' new suit in development for the Exploration Extravehicular Activity Services (xEVAS) contract. To accomplish these goals, our team will continue to advance Leto™, focusing on several major areas of development: a) rapid deployment for test environments, b) building digital twins for synthetic data generation, c) automated anomaly detection and prognostic machine learning model generation, and d) integration of generative AI. Each of these areas of development are described in greater detail in the following subsections.

A. Leto™ as a Test Environment

To assist with next generation spaceflight hardware development and begin valuable product data capture early in the development life cycle, our team is implementing a product-agnostic telemetry capture and analysis system. This testbed data capture system is designed to connect to any hardware test setup, extract telemetry in real-time, archive the data, and stream the data into a specialized branch of Leto™ specifically designed for analyzing test data. This system allows engineers to monitor tests in real time through statistical analysis and data visualization, establish expected telemetry bounds and receive alerts upon breaches, perform rapid diagnostics, and establish nominal distributions prior to actual certified product deployment; this method adds value to test engineers while enabling day-one, trusted spaceflight analytics and AI in the broader Leto™ spaceflight intelligence system.

B. Digital Twins

In addition to leveraging test data for day-one analytics, our team is also in the process of building digital twins for each product Leto™ monitors. The digital twins will help address data challenges presented in Section II, specifically they will help enable creation of algorithms for sensor drift detection, anomaly detection, and prognostics for new products with little-to-no spaceflight data. The digital twins, therefore, will be used to generate physics-informed synthetic data that should offer a solid foundational training set for Leto™ to learn from.

Relatedly, the digital twins will allow our team to simulate specific failure modes across all elements of a product. One of the challenges mentioned in Section II, and exemplified in Section V, is the lack of historic data for real failure events, despite having decades of historic data. The few failures experienced, evolving hardware configurations, and inconsistent modes of operation create difficulties in the data sets to be used for development of prognostics. Without ample representation of what failure events look like within the data, it is difficult to train an AI to predict failures. Therefore, to overcome this *rare-event* problem, the digital twins will be able to generate any desired failure mode such that the prognostics models will have enough events to train on. Additionally, many popular prognostics ML models rely heavily upon how long it has been since the last failure event when predicting remaining product life. This, however, creates a high-risk situation in which models often ignore the actual sensor telemetry as they are heavily weighted towards this time feature. Consequently, a hardware problem experienced shortly after a component has been recently replaced may not change the prognostic's prediction of remaining lifespan. By creating a large enough event dataset, our team will be able to better train its prognostics to focus on hardware telemetry and remove "time-since-last-event" bias.

The last goal of the digital twin effort is to build the mathematical framework for our team to pursue physics-informed machine learning (PIML). Statistical machine learning models are agnostic to real-world constraints, and therefore learn approximations through what is seen in the training data set. However, PIML constrains the models to known physics, therefore enhancing the models' performance as the underlying mathematics are both guided and unable to make predictions outside the realm of physical possibility.

C. Automated AIML Anomaly/Prognostics

As mentioned in Section I and II, one of the biggest challenges facing the spaceflight analytics domain is the dichotomy between the quantity of humans needed for monitoring spacecraft systems compared to the vast quantity of hardware systems, subsystems, and sensors to be monitored. As the commercial spaceflight industry accelerates, this dichotomy will only worsen without changing tools and methods. Additionally, advanced capabilities to address this dichotomy, such as using AI for system monitoring, takes significant time, effort, and domain specialization to develop that is not commonly found in the current pool of spaceflight experts. Therefore, our team is developing automated anomaly detection and prognostic capabilities for rapid model generation and deployment. The Leto™ data and software architecture has been designed to be highly standardized and consequently highly scalable. This standardization, along with the nature of all anomaly and prognostics problems being time-series of physics-based telemetry, allows for a single meta-model (i.e., a model that can idiosyncratically build and optimize both the data preprocessing pipelines and ML models themselves) to be built for any anomaly/prognostics problem.

D. Integration of Generative AI

In Section II our team touched on the challenge of SME knowledge loss through attrition. This is a very real problem across the industry, magnified by the long lifecycles of spaceflight hardware. Our team is looking to alleviate this issue through the integration of our vSME into Leto™ as a core capability, adding to the list of standard capabilities presented in Section IV. Having developed earlier prototypes and leveraging recent advancements in large language models (LLMs), our team is looking to integrate generative AI into the Leto™ spaceflight intelligence system, coupling generative AI with quantitative system monitoring to assist with hardware oversight.

VII. Conclusion

In this paper our team touched on a number of the challenges facing the spaceflight oversight and analytics domains, presented a vision for an intelligent system to address them, introduced Leto™ our spaceflight intelligence system developed with this vision in mind, and exemplified some of the key capabilities that Leto™ offers through a case study using the system to monitor the ISS WPA's Multifiltration Bed.

Specific to the examples, our team was pleased to see that our MFB remaining useful life model (predicting the first ionic breakthrough of the MFB) was accurate to within ten days, two months prior to the event, despite the lifetime data set including configuration and material changes in 2019 and 2020, respectively. With additional data, either through flight experience or synthetically generated with digital twins, our team looks forward to using Leto™ to predict the third ionic breakthrough to enable accurate, trusted removal and replacement planning that ensures full utilization of the MFB capacity while also protecting the Catalytic Reactor downstream. This is just one example of how Leto™ can be used to plan maintenance that maximizes hardware capacity, protects subsystems, and offers efficient use of crew time. We also demonstrated that with multi-model explainable anomaly detection built into Leto™, unexpected changes in performance can be detected, directing engineering to possible issues before they become problematic events. The anomaly detection functions of Leto™ will save time and minimize the risk that changes in system performance will go unnoticed for any reason, be it unintended human error or SME attrition.

As humanity's appetite and ambition to explore space grows, the current workforce will become strained and existing mission support frameworks may no longer work. Our team's aim is to alleviate this strain and enhance mission support through the capabilities that Leto™ offers: automated data collection and curation; system health and performance analytics, integrated artificial intelligence for prediction of failures and anomaly detection; and codification of subject matter expert knowledge enabling the addition of virtual subject matter experts to the existing pool of human capital. While there is still more work to do, the Leto™ spaceflight intelligence system has proven to be effective at resolving many of the stated issues in Section II and has proven to be highly scalable. Our team will continue evolving Leto™ to support a growing number of space destinations and address future mission needs as latency and distance from Earth drives greater autonomy requirements into mission architectures.

Acknowledgments

Funding for this work was solely provided by Collins Aerospace for independent research and development. The authors want to thank the Collins ISS Engineers and SMEs for their guidance and insights, which helped make this research and development investigation a success. Furthermore, the team would like to thank the ICES chairs for their thorough and thoughtful review of this paper prior to publishing.

References

[1]Space Foundation Editorial Team. "Space Foundation Releases The Space Report 2023 Q2, Showing Annual Growth of Global Space Economy to $546B." Space Foundation. July 25, 2023

[2]E. Bowman, L. Carter, J. Carpenter, N. Orozco, N. Weir, and M. Wilson, "Updated Performance Evaluation of the ISS Water Processor Multifiltration Beds," International Conference on Environmental Systems, 2014

[3]J. Williamson, J. Wilson, K. Robinson, and H. Luong, "Status of ISS Water Management and Recovery," International Conference on Environmental Systems, 2023