January 2021

NASA DATA STRATEGY

TABLE OF CONTENTS

LETTER FROM THE CHIEF DATA OFFICER

NASA recognizes that data is a strategic asset that must be operationalized to provide effective support that, combined with a network of partnerships, sustains future mission success in efforts to explore and discover our physical world.

An improvement to data management will enhance NASA's ability to lead efforts of exploration and discovery in an era of great competition. Effectively leveraging data will enable leaders, operators, scientists, and engineers to harness the power of data to capitalize on strategic opportunities and make informed, data-driven decisions. Machine learning and other advanced analytical techniques provide opportunities to create operational efficiencies and cost savings. Open and transparent data fosters innovative research and bolsters trust in agency operations. Carrying out these functions requires accurate, useful, timely, and accessible data. Embracing new data-driven concepts and leveraging data-backed innovations will improve our ability to learn from past missions, lead current technological advancements, and drive future exploration.

To enable this change, NASA is adopting leading data policies, standards, and capabilities as part of the Agency's overall Digital Transformation program. From enterprise governance models and data architectures to enhances in ML and AI capabilities, NASA is embracing a holistic organization-wide data transformation effort. This strategy builds on agency successes, evaluates lessons learned from past missions, and leverages momentum and opportunity from recent technological advancements. It addresses short term challenges identified across the agency and charts the course for establishing NASA as a data linchpin in both the private and public sectors. Our ability to continue to lead in efforts of exploration and discovery will require that we become leaders in operationalizing, protecting and effectively sharing our data resources at scale and at speed.

The NASA Data Strategy supports the Digital Transformation program by providing the Agency with a north star for the vision, mission, guiding principles, goals, and objectives necessary to navigate NASA's transformation into a data-centric enterprise. This strategy focuses on improvements to NASA's work, workforce, and workplace through enhanced data policies, personnel data skills, and data capabilities. The NASA Data Strategy applies to all of NASA's people and partners. It is everyone's responsibility to treat data as an asset. It must be managed, secured, and used to deliver on mission success and operational excellence. In this way, our Data Strategy will directly support the success of NASA's ability to continue to lead in missions of exploring our Planet Earth, our Solar System and beyond.

Ron Thompson

Chief Data Officer (CDO) & Deputy Digital Transformation Officer
National Aeronautics and Space Administration

EXECUTIVE SUMMARY

The National Aeronautics and Space Administration's (NASA) mission focuses on exploring and expanding our knowledge of Planet Earth and the Universe for the benefit of humanity. While this mission sits at the agency's core, there is not yet a fully cohesive data-centered approach to supporting the agency's goals of exploration and discovery.

Data is a key foundational pillar in digital transformation across the agency. It is an asset that should be leveraged to build on past achievements and chart the course for future mission successes. Data is increasingly considered the new currency for strategic decision-making. NASA recognizes that a holistic, organization-wide data strategy is necessary to effectively meet mission objectives and establish agency priorities. Such a strategy will allow for NASA to launch into a new age of data-driven innovation, technological advancement, and radical disruption to the status-quo. With a strong data policy as a foundation, NASA will be able to derive information from data, extract knowledge from information, and develop insights from knowledge – insights that can be used to inform policy and advance mission discoveries.

A cohesive agency-wide data strategy will push NASA to respond to current and future data needs. This Data Strategy describes a vision for being a data-driven organization. To achieve this vision, the agency must understand, access, and connect the agency's data; apply modern data science tools and techniques; cultivate data talent and skills among all NASA staff; and establish effective agency-wide data governance, data management, and data policy oversight.

The Federal Data Strategy has put federal agencies on notice that data and evidence need to play a more prominent role in how the federal government functions. NASA named its first Chief Data Officer (CDO), identified existing initiatives to serve as the foundation for agency-wide data efforts, and implemented various data initiatives under that broader federal mandate. This includes, but is not limited to, the creation of an agency-wide Data Governance Board (DGB), the development of an initial data strategy in some mission support areas, and the establishment of an agency data-standards working group. It now looks to build upon that preliminary progress and achieve larger organizational milestones as outlined in this Data Strategy document.

NASA's Data Strategy describes a vision, establishes agency level strategic goals for advancing data capabilities across the organization, and outlines capabilities and outcomes for those goals. NASA's vision is to *unleash the full power of data to accelerate NASA's ability to execute its missions and expand knowledge of the universe*. Goals to strengthen the data governance policies and procedures and advance a data-driven culture will enable NASA to build a foundation of data governance components and encourage a workforce that adheres to that governance model. A data-skilled staff working within the parameters of robust data policies and procedures will help to advance a one-data culture and help them harness the power of NASA's data and use it to their strategic advantage. Goals to *improve the agency's data principles* and *build a consolidated enterprise data capability* will enable NASA to focus on efforts to build, share and use data tools and analytic models for better decision-making, meet data transparency mandates, and encourage more research and innovation. These goals will

support the agency in its data mission: *elevate NASA's ability to leverage its data to explore, collaborate, learn, and share information within the agency, with its partners, and with the public.*

The NASA data strategy is intended to be revisited on a regular basis to balance the need for continuity and achievement as data and data-policies evolve over time. In developing these Data Strategy goals, NASA considered current needs and anticipated future business drivers, legislative mandates, and other environmental factors to ensure optimal alignment and successful execution. This Data Strategy must proactively keep NASA on the leading edge of data management and data science principles. To be a leader in this space, NASA will require constant attention and agility in the management of its data. This flexibility will allow the agency to recognize and embrace new data best-practices, advance its data science capabilities, and capitalize on opportunities for greater data interoperability across the organization to better serve all stakeholders. With progress towards this end, NASA will be well positioned to achieve its data vision to unleash the full power of its data and accomplish future mission success in its continued efforts of exploration and discovery.
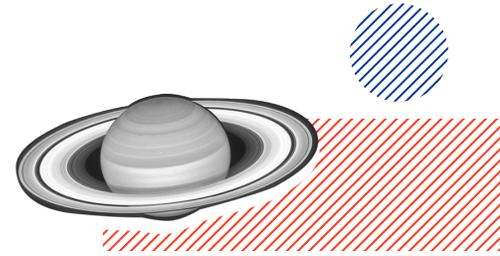
1. INTRODUCTION

The mission of NASA focuses on expanding the knowledge of our planet and our universe for the benefit of humanity. NASA seeks to accelerate exploration, collaboration, and learning with partners across the private and public domains. NASA leaders require data-driven insights to enable informed decisions, understand the current state of the Agency's operations and management, anticipate the organization's needs and opportunities, and to ultimately chart the path for future missions.

NASA recognizes that the rapid, global emergence of increasingly capable and connected digital technologies is impacting the Agency's missions, the expectations of its people, and what is needed for internal and external stakeholders to achieve their goals. In response, NASA has initiated an **Enterprise Digital Transformation (DT) Program** to lean into new technological advances and continue the agency's evolution into a data-centric organization.

An Enterprise DT Program will structure the Agency's efforts to transform NASA's:

1. **Work** – Deliver increasingly complex missions while leveraging increasingly complicated partnerships, on shorter timelines to achieve bolder outcomes that inspire the world.

2. **Workforce** – Create a seamless, integrated, and inclusive employee experience that energizes its people by fostering connection to the NASA enterprise, continuous skill growth, and pride in the Agency's ability to rapidly deliver high-value work.

3. **Workplace** – Optimize a 21st Century cyber-physical work environment that powers flexible, adaptable, efficient, and effective employee & partner teaming.

The transformations across NASA's work, workforce and workplace will require the agency to take advantage of data products from the past, present and future. As such, a specific level of ***hindsight, insight, and foresight*** will be utilized to ensure NASA is able to apply lessons learned from the past to the present and leverage all available information to anticipate mission outcomes, reveal opportunities, enable continuous improvement, and drive future mission successes.

FIGURE 1

HINDSIGHT	INSIGHT	FORESIGHT
Leverage the data that already exists. Analyze existing data sets through audits to better understand past operations and missions.	Exercise a necessary level of situational awareness to understand the present and act on current mission objectives.	Apply predictive analytics and modeling techniques to anticipate future outcomes to ensure future mission success.

**Work, Workforce, Workplace Transformation**

1.1 The Role of Data in Enterprise Digital Transformation

NASA Program Management created a framework for digital transformation anchored around a set of identified **Strategic Thrusts (ST)**. STs are shared enterprise-level digital enablers that the Agency can use as catalysts to accelerate the enterprise transformations noted above. Six primary STs are identified as the core catalysts for NASA's digital transformation:

1. **Process Transformation (PTx):** Transform products and processes to maximize efficiency and effectiveness to enable bolder missions faster.

2. **Artificial Intelligence/Machine Learning (AI/ML):** Harness machine capabilities to augment human intelligence in an era of big data.

3. **Model Based Everything (MBx):** Employ digital models to enable our people to address increasing complexity, scope, speed, uncertainty & changes.

4. **Data:** Ensure the data we need is Findable, Accessible, Interoperable, Reusable, Understandable, Secure, and Trustworthy (FAIRUST) to power data-driven decision making.

5. **Collaboration:** Enable agile teaming via seamless, secure internal and external collaboration

6. **Culture & Workforce:** Foster digital savvy, enterprise connection, and growth mindsets across the Agency.

FIGURE 2
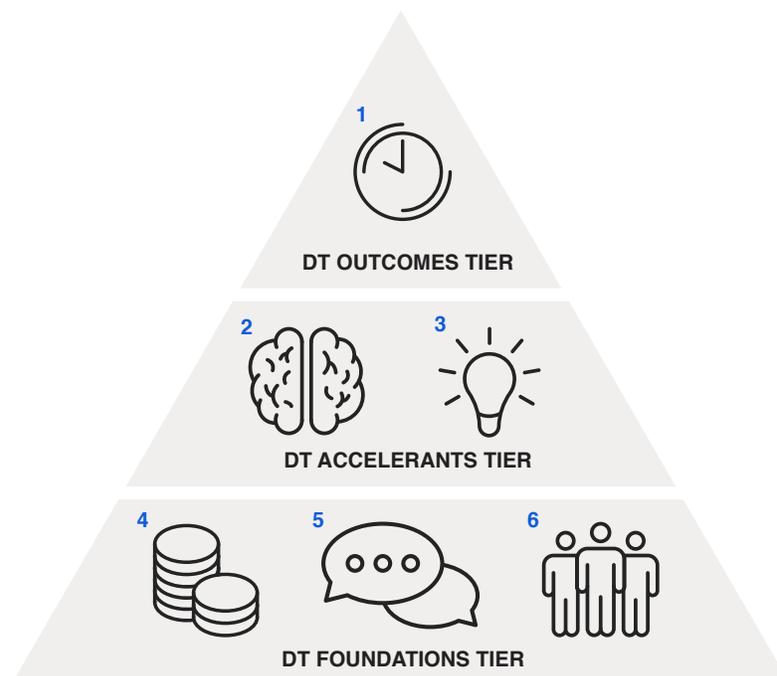
As illustrated in **Figure 2**, data serves as a foundational component of NASA's ability to achieve its core mission. Data underlies the success of all NASA activities, from effectively planning and executing missions, to understanding agency costs, determining improvements to facilities, training its workforce, and more. Data is not only important for achieving mission goals *within the Agency*, but also

with its partners and *with the public* to accelerate innovation, exploration, collaboration, and learning.

Given the central role of data in NASA's current missions and ongoing enterprise transformation initiatives, this document focuses on the Data ST and outlines a **NASA Data Strategy.** This document will serve as a north star that NASA will use to realize the full value of the Agency's data by providing a clear set of goals and objectives to guide the agency towards a data-driven culture that manages data as a strategic asset. The document also defines a roadmap for the Data ST, which includes the establishment of an **Enterprise Data Platform (EDP)** within NASA.

1.2 Problem Statement

Data is a crucial facet of NASA mission oriented and operational processes. NASA historically managed data platforms across its centers and facilities in a decentralized manner. Independently deployed functions across the enterprise have led to disjointed and duplicative work efforts, tools, infrastructure, and resources. This disconnected data infrastructure inhibits NASA's ability to leverage its rich array of operational, scientific and research data products.

While this fragmented approach supports the day-to-day operations of individual business units, it creates limitations in achieving agency-level data enterprise and digital transformation goals. While some areas of the Agency, such as NASA science archives, have adopted modern strategies for data management, much of the Agency's data remains siloed and inaccessible. The lack of interoperability between Agency data systems makes it increasingly difficult for both internal and external stakeholders to effectively discover, understand, integrate, access, and use valuable scientific and business data sources. Additionally, the Agency lacks a consistent methodology to understand, tag, and categorize its data sources which, in turn, compromises its ability to enhance an enterprise security posture and enforce a mature data governance framework.

The primary issues seen in the Agency's data pipeline today are summarized below.

NO COMPREHENSIVE SEARCH	SILOED DATA SETS	LOCAL ANALYTICS ENVIRONMENTS	DECENTRALIZED VISUALIZATION TOOLS	NON-SHARABLE DATA ASSETS
Steep learning curve to understand the data landscape throughout NASA, which leads to misuse of data and lack of collaboration across centers	Data (e.g. Science, Engineering, Flight, Safety, Financial, etc.) scattered across independently managed systems leads to redundancy and confusion among end users and inconsistent reporting and analytical results	Exporting data for analysis is limiting and time consuming as data analysts spend large portion of time on data cleansing tasks with limited resources	Multiple independent instances of the same visualization tools throughout the enterprise makes it difficult to manage reports and create standard definitions and templates for leadership	Without a centralized data asset catalog, analysts lack the necessary tools for optimal collaboration and validation

Along with implementing transformational data architecture changes, NASA must also make strong progress shifting its data culture from legacy "knowledge is power" approaches to data best practices, such as establishing trusted and consistent data sources, sharing data across business units, and using data to solve business challenges. The Agency must address how to positively shift employee attitudes and perception along with targeted workforce development and training. Modern skills, culture, and behaviors will enable NASA to take full advantage of any new data-related capabilities.

A modern data governance model and a consolidated data architecture will create an environment where NASA can build trust and get more value from its data, effectively set enterprise priorities, achieve consistent results across missions and enable better use and impact of its data, both within NASA and externally.

1.3   Scope of Strategy

To be successful in its goal of enterprise data transformation, NASA must take a multi-pronged approach to implementing this Data Strategy. The focus areas for this approach will cut across people, processes, technologies, policies, culture, skills, organization, and governance.

The NASA Data Strategy applies to the entire NASA Agency and its data, on whichever systems that information resides. The strategy is consistent with and uses guidance outlined in the Federal Data Strategy to align the strategic direction and underlying principles of NASA's goals to those being adopted across the Federal Government.

Figure 3 summarizes the NASA Data Strategy. It describes a vision, establishes agency level strategic goals for advancing data capabilities across the organization, and outlines capabilities and outcomes for those goals. NASA's vision is to *unleash the full power of data to accelerate NASA's ability to execute its missions and expand knowledge of the universe*. This agency-wide effort will require discussions across the organization about data priorities that will help improve data maturity and advance NASA's ability to leverage data, operationalize and optimize data governance, and drive culture change for the benefit of its internal and external stakeholders.

FIGURE 3

NASA DATA STRATEGY

Vision: Unleash the full power of data to accelerate NASA's ability to execute its missions and expand knowledge of the universe.

Mission: Elevate NASA's ability to leverage its data to explore, collaborate, learn, and share data and information within the Agency, with its partners and with the public.

GOALS

01
STRENGTHEN
Agency-wide data governance policies and procedures

02
ADVANCE
A data-driven culture across the agency

03
IMPROVE
Agency data products' adherence to FAIRUST data principles

04
BUILD
An Enterprise Data Architecture to provide a consolidated data capability

OBJECTIVES

- Empower the Chief Data Officer and establish agency-level data governance boards to continue to create and promote data policy.

- Create a centralized security architecture and data governance framework to manage core metadata on all systems.

- Identify data stewards for all NASA data and systems.

- Maintain system level data management plans and designated POCs who are responsible for the curation of data and content in their systems.

- Embrace a 'one-data' culture free of siloes.

- Form inter-agency communities of interest to share analyses and best practices.

- Identify and invest in data-related skills and training, professional development, retention and recruitment.

- Shift from a legacy mindset of restricting information on a need-to-know basis to a mindset where appropriate sharing is rewarded.

- Ensure data is Findable, Accessible, Interoperable, Reusable, Understandable, Secure, Trustworthy.

- Create and store data artifacts that can be reused and shared across the enterprise.

- Deploy programmatic tools from which employees, partners, and the public can access all data.

- Move to modern, zero-trust technologies to better secure data at the source and enhance consumer's trust in NASA data.

- Consolidate data and standardize ingestion pipeline to allow for analysis from a single Data Platform.

- Create reusable data assets that are stored and analyzed within an Enterprise Data Platform.

- Introduce catalog and search capabilities for finding data assets to pull into agnostic end user toolsets.

- Centralize BI and visualization toolsets by standardizing an enterprise suite of analytics products.
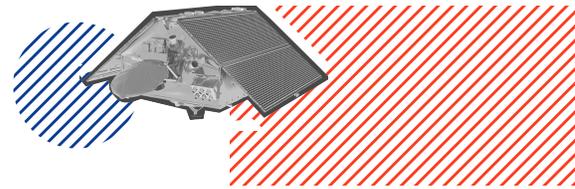
GUIDING PRINCIPLES

Ethical Governance   *Conscious Design*   *Learning Culture*

Goals to *strengthen the data governance policies and procedures* and *advance a data-driven culture* will enable NASA to build a foundation of data governance components and encourage a workforce that adheres to that governance model. A skilled staff working within the parameters of robust data policies and within a one-data culture and will help them harness the power of NASA's data and use it to their strategic advantage in operational and mission support.

Goals to *improve the agency's data products' adherence to data principles* and *build a consolidated enterprise data architecture capability* will enable NASA to focus on efforts to build capabilities and use data for better decision-making, meet data transparency mandates, and encourage more research and innovation. These goals will contribute to the enhancement of program operations and support the agency in its data mission: *elevate NASA's ability to leverage its data to explore, collaborate, learn, and share information within the agency, with its partners, and with the public.*

2.0 VISION, MISSION, AND GUIDING PRINCIPLES

NASA's vision and mission regarding data management reflect the overall vision and mission of the agency, which are anchored around the themes of exploration, discovery, and sharing information. To realize the full potential of data, NASA Leadership's continued commitment is critical to transforming the enterprise and employing new capabilities, where data drives policies, operations, and missions.

2.1 Vision & Mission

A core tenet of the NASA Data Strategy is the understanding that data is an asset and serves as an integral part of the business operations and missions it supports. It is critical that data be of the highest quality, accurate, complete, protected, and trustworthy to enable the data consumers to achieve their mission goals.

FIGURE 4

Vision: Unleash the full power of data to accelerate NASA's ability to execute its missions and expand knowledge of the universe.

Mission: Elevate NASA's ability to leverage its data to explore, collaborate, learn, and share data and information within the Agency, with its partners and with the public.

The concept of continuous improvement is embedded in the vision and mission of NASA's data strategy. A primary outcome of the strategy is for the enterprise to establish a data-driven culture that encourages and enables the sharing of data within the organization, with partners and with the public. A data-centric culture will create an environment that promotes the development and growth of a positive feedback loop, as illustrated in **Figure 5**, and adopts a "one data" mindset. This loop will be a continuous cycle in which data policies, data management, and data product scalability further builds an organizational culture that uses, shares, and trusts its data

FIGURE 5
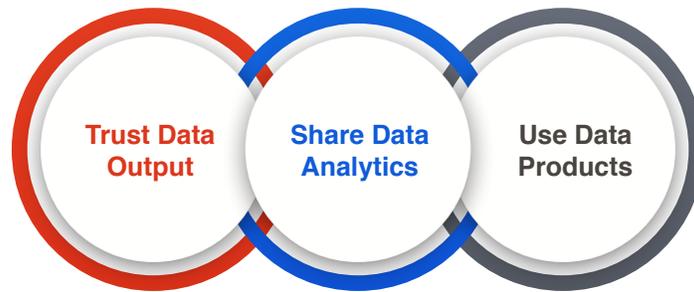
2.2 Guiding Principles

NASA is fully committed to adopting the guiding principles established in the **Federal Data Strategy**[1]. The guiding principles of *ethical governance, conscious design*, and strong *learning culture* underpin all goals and objectives outlined in this Data Strategy. The Agency abides by these values and intends to continually self-assess to ensure on-going alignment with all the activities of the Data Strategy.

ETHICAL GOVERNANCE

- **Uphold Ethics:** Monitor and assess the implications of federal data practices for the public. Design checks and balances to product and serve the public good.

- **Exercise Responsibility:** Practice effective data stewardship and governance. Employ sound data security practices, protect individual privacy, maintain promised confidentiality, and ensure appropriate access and use.

- **Promote Transparency:** Articulate the purposes and uses of federal data to engender public trust. Comprehensively document processes and products to inform data providers and users.

Conscious Design

- **Ensure Relevance:** Protect the quality and integrity of the data. Validate the data are appropriate, accurate, objective, accessible, useful, understandable, and timely.

- **Harness Existing Data:** Identify data needs to inform priority research and policy questions; reuse data if possible and acquire additional data if needed.

- **Anticipate Future Uses:** Create data thoughtfully, considering fitness for use by others; plan for reuse and build interoperability and data standards from the start.

- **Demonstrate Responsiveness:** Improve data collection, analysis, and dissemination with ongoing input from users and stakeholders. The feedback process is cyclical; establish a baseline, gain support, collaborate, and refine continuously.

---

1   Vought, Russell T. Principles for Leveraging Data as a Strategic Asset. Office of Management and Budget (OMB). strategy.data.gov/assets/docs/federal-data-strategy-principles.pdf. Accessed 8 Jan 2021.

> ### LEARNING CULTURE
>
> - **Invest in Learning:** Promote a culture of continuous collaborative learning with and about data through ongoing investment in data infrastructure and human resources.
> - **Develop Data Leaders:** Cultivate data leadership at all levels of the workforce by investing in training and development about the value of data for mission, service, and the public good.
> - **Practice Accountability:** Assign responsibility, audit data practices, document and learn from results, and make needed changes.

3.0 GOALS AND ENABLING OBJECTIVES

The goals of this Data Strategy address new and emerging themes as defined in the Federal Data Strategy, such as open data by default, inter-agency data sharing, and data standardization. Improving strategic data management will help to solve persistent data challenges and offer opportunities to close capability gaps to allow the Agency's workforce to make data-backed decisions in operational tasks and mission assignments. These goals are broad and represent core components of the NASA data strategy. The four interconnected goals are outlined in **Figure 6** and are comprised of measurable objectives, each defined to be meaningfully achieved in a reasonable timeline.

FIGURE 6

DATA STRATEGY GOALS

01 STRENGTHEN
Agency-wide data governance policies and procedures

02 ADVANCE
A data-driven culture across the agency

03 IMPROVE
Agency data products' adherence to FAIRUST data principles

04 BUILD
An Enterprise Data Architecture to provide a consolidated data capability

3.1 Goal 1: *Strengthen Data Governance*

Strengthening data governance across the NASA enterprise serves to improve the framework by which data is captured, stored, and managed across the agency. The future state objectives for this goal are to:

- Empower the Chief Data Officer (CDO) and establish agency-level data governance boards to continue to create and promote data policy.

- Create a centralized enterprise security architecture and data governance framework to manage core metadata on all NASA system.

- Identify data stewards for all NASA data and systems.

- Maintain system level data management plans and designated POCs who are responsible for the curation of data and content in their systems.

3.2 Goal 2: *Advance a Data-Driven Culture*

Advancing a data-driven culture across NASA serves to fundamentally transform the way in which employees' view, use and promote the agency's data. A transformation in the agency's culture will ultimately be defined by the following a human-centered approach to achieve future state objectives:

- Embrace a 'one-data' culture free of siloes.

- Form inter-agency communities of interest to share analyses and best practices.

- Identify and invest in people development focused on data-related skills and training, professional development, retention, and recruitment.

- Shift from a legacy mindset of restricting information on a need-to-know basis to a mindset where appropriate sharing is encouraged and enabled. Shared data assets will improve the quality and access of data products across the agency, streamline data analysis and encourage future innovation.

3.3 Goal 3: *Improve Data Products & Data Principles*

Improving data products and data principles across the Agency will ensure that NASA-owned data meets a threshold of consistency. This consistency will increase the probability it is used, trusted, and shared across the enterprise, with partners, and with the public. A future state where data products and data standards adhere to standard principles will allow the agency to:

- Ensure NASA's data products adhere to the guiding principles of FAIRUST, in which data is findable, accessible, interoperable, reusable, understandable, secure, and trustworthy. FAIRUST principles are defined in **Figure 7**.

- Create and store data artifacts that can be reused and shared across the agency.

- Deploy programmatic tools from which employees, partners, and the public can access all data.

FIGURE 7

WHAT IS FAIRUST?

FINDABLE

Consumers can find the needed data.

ACCESSIBLE

Consumers can retrieve the data.

INTEROPERABLE

Consumers and producers have a common representation and comprehension of data.

REUSABLE

Consumers can leverage existing data products to feed decisions as opposed to reinventing the data set.

UNDERSTANDABLE

Consumers can find descriptions of data to recognize the content, context and applicability.

SECURE

Consumers know that data is protected from unauthorized use and manipulation.

TRUSTWORTHY

Consumers can be confident that data can be used for decision making.

- Move to modern, zero-trust technologies to better secure data at the source and enhance consumer's trust in NASA data.

3.4 Goal 4: *Build an Enterprise Data Architecture*
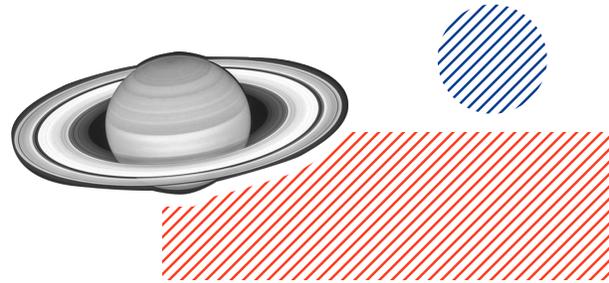
Building an Enterprise Data Architecture at NASA serves to provide a holistic solution to the data strategy and the data needs of the NASA enterprise. The future state objectives for an improved Enterprise Data Architecture will allow NASA to:

- Build and implement a robust, flexible Enterprise Data Platform (EDP) that serves the entire agency.

- Consolidate data and standardize ingestion pipelines to allow for analysis from a single platform.

- Create reusable data assets that are stored and analyzed within the platform.

- Introduce catalog and search capabilities for finding data assets to pull into agnostic end user toolsets.

- Centralize BI and visualization toolsets by standardizing an enterprise suite of analytics products.
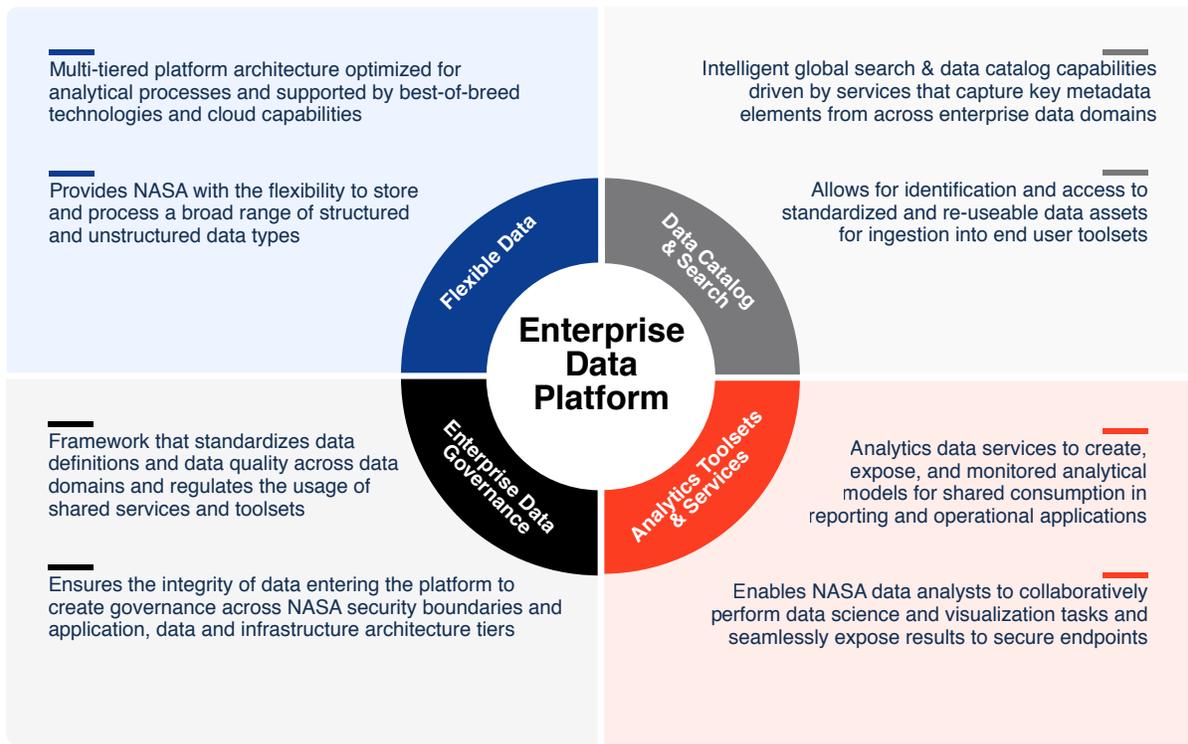
An EDP will provide NASA with a foundational enterprise data architecture that serves as an enabling self-service solution with seamless data management and data science capabilities for the Agency to achieve its defined Data Strategy goals. An EDP will provide NASA with a consolidated data platform that brings data analysis and visualization out of disparate applications into a shareable repository. The platform will be supported by automated services and a comprehensive toolset to create:

1. A searchable catalog of data assets (tables, result sets, models, etc.),

2. Governed data to create reusable and understandable data products, and

3. A simplified and controlled security framework to manage access more consistently and effectively.

A comprehensive EDP architecture will directly support the NASA data consumer by allowing users to create and share data using simple, integrated, self-service data product workflows efficiently and collaboratively. Data professionals will be able to derive and convey insights to support timely and sound decisions for leadership and other stakeholder audiences. An intelligent search service will allow consumers to find and recall data products or other relevant information within their authorized access limits, allowing for quick, efficient, and trustworthy data gathering techniques.

Figure 8 summarizes the key benefits provided by an Enterprise Data Platform.

FIGURE 8

Multi-tiered platform architecture optimized for analytical processes and supported by best-of-breed technologies and cloud capabilities

Provides NASA with the flexibility to store and process a broad range of structured and unstructured data types

Intelligent global search & data catalog capabilities driven by services that capture key metadata elements from across enterprise data domains

Allows for identification and access to standardized and re-useable data assets for ingestion into end user toolsets

Framework that standardizes data definitions and data quality across data domains and regulates the usage of shared services and toolsets

Ensures the integrity of data entering the platform to create governance across NASA security boundaries and application, data and infrastructure architecture tiers

Analytics data services to create, expose, and monitored analytical models for shared consumption in reporting and operational applications

Enables NASA data analysts to collaboratively perform data science and visualization tasks and seamlessly expose results to secure endpoints

Flexible Data
Data Catalog & Search
Enterprise Data Governance
Analytics Toolsets & Services

Enterprise Data Platform

4.0 DATA DEVELOPMENT TRACKS

NASA utilized the **Federal Data Maturity Model** to identify gaps in needed capabilities at the agency. Based on this model, 6 parallel development tracks were established to focus on Agency activities that will drive progress towards the NASA Data Strategy Goals. The development tracks are *1. Data Governance; 2. Data Process & Policy; 3. Data Workforce; 4. Data Culture; 5. Data Tooling; and 6. Data Management and Architecture.*

Each of these development tracks support one or more Strategic Goals. The accomplishments made within each track will directly drive progress towards the goals to which they support.

Track 01: DATA GOVERNANCE

Establish governing structures to enforce data-related policies and common processes, and enhance the communication, coordination, and cooperation between data experts.

GOALS SUPPORTED:

01
Strengthen Agency-wide data governance policies and procedures

04
Build an Enterprise Data Architecture to provide a consolidated data capability

Track 02: DATA PROCESS & POLICY

Leverage the Data Governance Board and create Agency-level working groups to execute the agency Data Strategy, analyze progress towards strategic goals and objectives, and coordinate the development of needed policy & technical efforts to ensure insights and analyses align across the Agency.

GOALS SUPPORTED:

01 Strengthen Agency-wide data governance policies and procedures

04 Build an Enterprise Data Architecture to provide a consolidated data capability

---

Track 03: DATA WORKFORCE

Address existing and future workforce needs by identifying required data-related skills and associated training, professional development, retention, and recruitment opportunities.

GOALS SUPPORTED:

02 Advance data-driven culture across the agency

04 Build an Enterprise Data to provide a consolidated data capability

---

Track 04: DATA CULTURE

Transform culture at the agency to promote concepts of data sharing, data stewardship and trust in data. Adopt mindsets of sharing data by default rather than protecting data to remove data limitations and duplication caused by siloes.

GOALS SUPPORTED:

02 Advance data-driven culture across the agency

04 Build an Enterprise Data Architecture to provide a consolidated data capability

---

Track 05: DATA TOOLING

Acquire and maintain shared tools that allow the workforce to better analyze, visualize, understand, create, merge, reuse, clean, create and manage data across NASA. Leverage these tools to support progress towards goals of the interdependent AI/ML Strategic Thrust.

GOALS SUPPORTED:

03 Improve Agency data products' adherence to FAIRUST data principles

04 Build an Enterprise Data Architecture to provide a consolidated data capability

---

Track 06: DATA MANAGEMENT & ARCHITECTURE

Develop the required technological capabilities, standards and management practices which will lead to the optimal curation and use of data through a FAIRUST data architecture with shared enterprise capabilities.

GOALS SUPPORTED:

03 Improve Agency data products' adherence to FAIRUST data principles

04 Build an Enterprise Data Architecture to provide a consolidated data capability

5.0 STRATEGIC ROADMAP

Operationalizing the vision, mission and principles of the NASA data strategy is an enduring, agency-wide effort. While there has been work accomplished to date across the strategic goals and development tracks, more near-term and long-term milestones are planned to drive the data strategy towards the desired future state.

5.1 Progress to Date

Progress has already been made by the Agency across the six development tracks in relation to NASA's data strategy goals. These efforts are ongoing and are made possible by effective cross-team collaboration across the enterprise.

Track 01: DATA GOVERNANCE

ACCOMPLISHMENTS TO DATE

- ✓ A Charter for Data Governance Boards has been developed and signed. This board has been operational in Spring 2020.

- ✓ The CDO role was defined and filled as of August 2019.

- ✓ Data Governance standards have been established in prioritized missions.

- ✓ Data management plans are defined and enforced for prioritized data systems.

Track 02: DATA PROCESS & POLICY

ACCOMPLISHMENTS TO DATE

- ✓ Independent work to develop data strategy in some mission and mission support areas. The SMD Data Strategy is an example.

- ✓ An Agency Data Standards Working Group has been established.

Track 03: DATA WORKFORCE

ACCOMPLISHMENTS TO DATE

- ✓ Decentralized trainings have occurred for target data skills at the Agency.

- ✓ Initial recruitment strategy for data professionals.

Track 04: DATA CULTURE

ACCOMPLISHMENTS TO DATE

- ✓ Key data cultural issues identifiedInitial recruitment strategy for data professionals.

Track 05: DATA TOOLING

✓ Defining tools strategy

✓ Created tiger team to understand Business Intelligence (BI) tooling

✓ Scattered sites linking work and data

✓ An agency-wide cloud-based environments capability has been developed for hosting and processing data (storage, processing, and database services)

Track 06: DATA MANAGEMENT & ARCHITECTURE

✓ Isolated data architectures and data standards across the Agency

✓ Lessons learned have been collected for practical implementation of data management at the Agency

✓ Isolated data dictionaries and metadata standards for many domains of data

5.2 Roadmap to Future State Goals

While progress has been made towards each primary data goal, there remains a series of milestones that NASA has identified that will accomplish near term, long-term, and future state objectives. This section outlines a roadmap for each of the agency's primary goals to 1. Strengthen Data Governance; 2. Advance a Data-Driven Culture; 3. Improve Data Products and Standards; and 4. Build an Enhanced Enterprise Data Architecture.

**01
STRENGTHEN**
Agency-wide data governance policies and procedures

To **strengthen data governance** at NASA, the agency must focus on improvements to existing data governance practices and updates to current data processes and policies. The roadmap to strengthen data governance across the agency encompass the strategic thrusts that include process improvements (data governance boards), technical improvements (data governance platforms, automated tagging, common metadata) and workforce improvements (development of data stewards).

Future State:
To reach a future state, the agency will need to utilize the data governance boards to put the needed changes in policy, access, and access controls to key data systems and approve needed metadata standards.

02
ADVANCE
A data-driven culture across the agency

To **advance a data-driven culture** at NASA, the agency must focus on developing a culture and a workforce that is data-centric in all activities. Addressing cultural issues related to data is paramount for success in all other development tracks. Efforts will be needed to identify, strategize, and execute initiatives to change how technology around data is used. A fully transformed data culture will allow for all NASA business units to have multidisciplinary teams of data professionals and business subject matter experts working together to solve analytical problems.

Future State:
To reach a future state, the agency must hire and train its workforce to support data activities. This will require the creation of specific data professional roles and an understanding of the skills to empower other data-related initiatives, such as data management, data curation and data preparation for use in advanced analytics.

03
IMPROVE
Agency data products' adherence to FAIRUST data principles

To **improve data standards** at NASA and align them to industry standards, the agency must focus on updates to data tooling and data products. Changes in data tools, products, and systems technology is highly dependent on the data governance and data policy outlined in Goal 1. A durable data governance model and sound data policies will allow for the formation of a robust data architecture which will empower the creation and use of more capable data tools. In the future states, NASA employees at all levels will be able to navigate this data platform to find, use and share data assets with little effort.

Future State:
To reach a future state, NASA will strive to align the agency data standards to qualities outlined by the FAIRUST model which seeks to make data findable, accessible, interoperable, reusable, understandable, secure, & trustworthy.
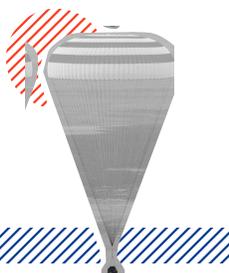
04
BUILD
An Enterprise Data Architecture to provide a consolidated data capability

To **build an enterprise data architecture** at NASA, the agency must focus on implementing advanced data solutions that promote collaboration and centralization of Agency data sources and tooling. This will provide an enterprise capability to advance all of NASA's data goals. An enhanced data architecture will be accomplished through the development, deployment, and maintenance of an enterprise data platform (EDP). The EDP will strengthen the agency's data governance policies, provide a platform to share data assets, advance a culture of collaboration, and improve data principles by enforcing common data standards across the enterprise.

Future State:

To reach a future state, NASA must develop and deploy an enterprise data platform. An initial platform will be built with a select set of priority data assets. Over time, iterative improvements to the architecture will allow for the expansion of data sources and the migration of more data products to the platform. These improvements will provide a platform that reinforces NASA's goal to transform into a data-centric organization where shared data assets can be used to improve outcomes for future missions.
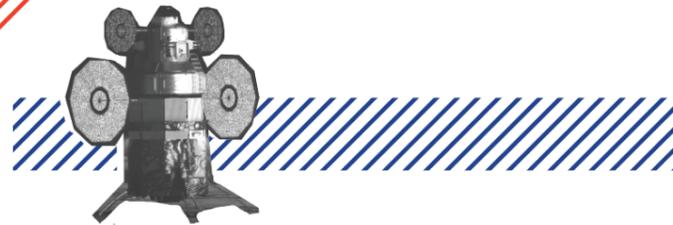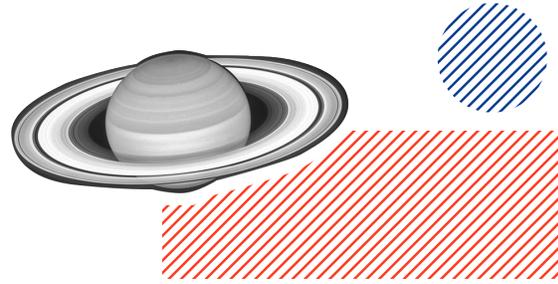
Figure 9 outlines the near term, long term, and future state milestones for each development track, which are supporting drivers for each of the primary Data Strategy goals focused on in this section and further described in Section 3.0.

FIGURE 9

		NEAR TERM OBJECTIVES	LONG TERM OBJECTIVES	FUTURE STATE OBJECTIVES
01 STRENGTHEN Agency-wide data governance policies & procedures	DATA GOVERNANCE	• Functioning Governance Boards • Data governance defined for core metadata. • Data governance platform available. • Automated data taggers for agency.	• Agency-wide governance platform. • Agency-wide curation of data governance for core metadata on key systems. • Data stewards for key systems.	• All data is governed through agency-level collaboration. • Data governance of metadata on all systems. • Data stewards for all NASA systems.
04 BUILD An Enterprise Data Architecture to provide a consolidated data capability	DATA PROCESS & POLICY	• Functioning agency level data strategy working groups. • Draft policy for data governance and default access, roles and mgmt plans for all systems.	• Robust planning for data. • Robust data policies in place. • Alignment with the Federal Data Strategy.	• Data policies and plans reflect data strategy. • Strategy & policy can rapidly pivot to respond to mission challenges.
02 ADVANCE A data-driven culture across the agency	DATA WORKFORCE	• Agency Data Personnel training. • Professional path created for Data Personnel. • Data professionals with SMEs. • Execute agency recruitment for Data Personnel.	• Data professionals integrated with SMEs in all business areas. • Some teams solving BI problems. • Agency workers can access data trainings.	• All areas of NASA have multidisciplinary teams of Data Professionals and Business SMEs solving analytical problems.
04 BUILD An Enterprise Data Architecture to provide a consolidated data capability	DATA CULTURE	• Some data analytics and quality processes are routine. • Many business data are managed across the agency and have some defined standards.	• High demand for data across the agency. • Data-centric principles drive decision making. • Business data are governed with cross-functional use in mind. Standards are uniform.	• NASA embraces a 'one-data' culture. • Inter-agency communities of interest share analyses and best practices.
03 IMPROVE Agency data products' adherence to FAIRUST data principles	DATA TOOLING	• Self-service means available to install BI tools • AI/ML storefront, Expert Portal in use. • Cloud based BI, AI/ML, Analytics and Dev tools. • Agency shares dashboards, containers & code.	• Analytics used to understand data are integrated with roadmap planning. • Core BI tools and self-service library are available. • Wide adoption of BI and Expert Portal.	• NASA employees at all levels can access and use needed data tools with little effort. • Data tooling applied to all data sources.
04 BUILD An Enterprise Data Architecture to provide a consolidated data capability	DATA MGMT. & ARCHITECTURE	• Agency-level data architecture and services provide basic data discovery and sharing. • Zero trust data available for some systems. • Agency data catalog available for OMB.	• FAIRUST data architecture in place. • Zero trust data widely available. • Only legacy data is siloed. • Key data can be programmatically accessed (APIs).	• All data adhere to data sharing policy. • Zero trust implemented for all data. • Programmatic access to all data at the agency.

5.3 Roadmap for an Enterprise Data Platform (EDP)

The creation of an enterprise data platform will advance each of NASA's four data goals in a significant way. Initially, it will become the basis for building an enhanced data architecture by introducing a modern enterprise data platform that serves all the Agency's data goals. Once implemented, this platform can then be continuously updated to meet NASA's evolving data needs. The EDP capability will provide NASA and its stakeholders with a streamlined platform to advance their data goals to **strengthen, advance,** and **improve** NASA's data posture.

More specifically, the EDP will provide NASA with a way to:

1. **Strengthen** data processes with a consolidated repository to track metadata and governance policies.

2. **Advance** the data workforce with a centralized platform to allow for a culture of collaboration across the agency.

3. **Improve** the available data tools with a developed toolset to complete more advanced analytics and data work.

There are many synergies between an EDP and NASA's data goals. Data Governance serves as a crucial capability for ensuring organizations can store, manage, and securely access its wide array of data sources in a scalable manner across the enterprise. The EDP will provide NASA with the required framework and tools to effectively govern and employ dense networks of data sets which are used across its range of applications and phases of data processing. These updates will enable for more complex data processing capabilities which are essential to NASA programs, operations, and strategic missions.

To ensure the advancement of NASA's goals, the creation of the EDP should not be completed by an isolated team. Instead, data stewards for all data systems will be involved to complete data inventory assessments. These assessments will be key inputs to the design of the EDP because they will help to identify priority data sets and use cases to be made available on the EDP. Data stewards will work within the initial data governance model and identify the agency's priority data sets. Their main goal will be to understand the three main principles (defined in **Figure 10**) around priority data sets to accurately catalog them.

FIGURE 10

Key Data Assessment Principles:

What is the data?

Who owns the data?

How is the data used?

Which data sets are considered priority by the agency?

Once priority data sets are identified and accurately cataloged, they can be used as assets in the broader EDP roadmap to launch. With data sets in hand, an initial EDP proof of concept can be developed, more data can be consolidated, and an enterprise capability can evolve to scale. As the enterprise capability matures, so too do the underlying components, including the platform architecture, the analytic toolset, the catalog and search functionalities, the data visualization suite,

and the overall data governance model. This roadmap for the EDP follows a trajectory that starts with an initial prototype and ends with a full-scale enterprise platform. A full roadmap is illustrated in **Figure 11**.

FIGURE 11

NOV 2020	SPRING 2021	FALL 2021	SPRING 2022
PROTOTYPE Demonstrate Capability	**CONSOLIDATE** Centralize Reporting & Data Catalog Tools	**TRANSFORM** Establish Integration Tools & Analytics Hub	**SCALE** Expand to All of NASA Enterprise
• **Prototype System** deployed on NASA development environment • **Data Storage** (AWS RDS) **Reporting** (Tableau) and **Data Ingestion** (Alteryx) tools installed • Ingestion, Storage, & **Visualization** processes established on sample datasets	• **Enterprise Tableau** and **Alteryx Data Catalog** implemented • **Consolidation of Tableau Sites** begins with ingestion of first data sets into EDP • **EDP Governance Board** established, with enterprise Data Stewards identified	• **Alteryx Analytics Hub** and **AWS Data Lake** implemented • **Transition of Data Analysis Culture** from NASA Center level to agency level data domains • **Agency Data Onboarding Plan** approved by Data Stewards & EDP Governance Board	• **Snowflake Data Platform** implemented to scale performance • **Analytical Modeling Suite** deployed to manage models and transition them to operations • **EDP Governance Playbook** established to standardize Agency governance processes

INCREASING EDP CAPABILITY MATURITY THROUGHOUT EXECUTION OF ROADMAP PHASES

NOV 2020	SPRING 2021	FALL 2021	SPRING 2022
DATA PLATFORM	DATA PLATFORM	DATA PLATFORM	DATA PLATFORM
ANALYTICS	ANALYTICS	ANALYTICS	ANALYTICS
CATALOG & SEARCH	CATALOG & SEARCH	CATALOG & SEARCH	CATALOG & SEARCH
DATA VISUALIZATION	DATA VISUALIZATION	DATA VISUALIZATION	DATA VISUALIZATION
DATA GOVERNANCE	DATA GOVERNANCE	DATA GOVERNANCE	DATA GOVERNANCE
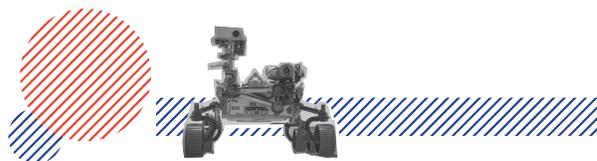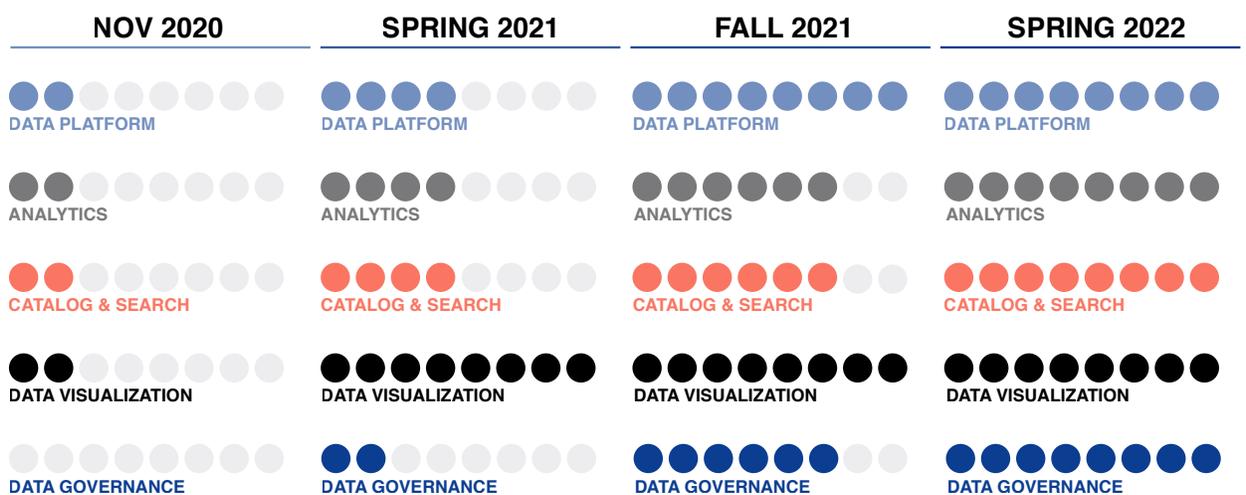
6.0 CONCLUSION

Data is increasingly the driver of every process and decision within the NASA Agency. The enterprise data strategy serves as an approach to transform NASA into a data-driven organization. This transformation requires effective and robust data management paired with a strong relationship between data gatherers, data analysts, data stewards and data users. Data will be treated as an asset of the highest caliber, and NASA leaders will govern, protect, and curate that data to ensure it is effectively and responsibly shared with internal and external stakeholders. NASA must cultivate a community of data professionals and equip all personnel with the data skills and tools necessary to preserve its advantage in the overall mission of exploration and discovery of our Earth, our solar system and beyond.