



# FY25-28

## ENTERPRISE DATA STRATEGY

U.S. Department of Energy

# Table of Contents

---

|                                                                                                                 |    |
|-----------------------------------------------------------------------------------------------------------------|----|
| Letters from DOE Leadership                                                                                     | 1  |
| Enterprise Data Strategy Executive Summary                                                                      | 4  |
| Vision                                                                                                          | 4  |
| Guiding Principles                                                                                              | 4  |
| Goals & Objectives                                                                                              | 4  |
| Introduction                                                                                                    | 5  |
| Enterprise Data Management at DOE                                                                               | 5  |
| Our Collective Approach to Creating the FY25-28 DOE Enterprise Data Strategy                                    | 7  |
| Our Path Forward: Implementing the Enterprise Data Strategy                                                     | 8  |
| Goals & Objectives                                                                                              | 9  |
| Goal 01: Establish & Scale Federated Data Governance & Stewardship                                              | 9  |
| Goal 02: Build & Maintain Data Architecture & Infrastructure Foundations                                        | 10 |
| Goal 03: Strengthen Capabilities that Improve Discoverability, Usability, Quality & Trustworthiness of Our Data | 11 |
| Goal 04: Foster & Sustain Our Next Generation Data & Analytics Workforce                                        | 12 |
| Goal 05: Ensure the Trusted & Scalable Use of Data for AI                                                       | 13 |



# Letter from The Deputy Secretary



I am pleased to present the inaugural U.S. Department of Energy (DOE) Enterprise Data Strategy. This Strategy is the product of a comprehensive and collective effort by representatives from across the Department to unlock the full potential of DOE's complex data landscape. Our goal is to better inform every aspect of decision-making at DOE to improve impacts and outcomes across our energy, science & innovation, safety & security, and operational missions. The development of a robust and forward-looking Data Strategy is especially crucial for DOE because we are the leading generator of both unclassified and classified scientific data through the world's largest collection of advanced experimental facilities.

DOE and its network of 34 national user facilities produce instruments for the broader scientific community to enable research while also generating tremendous volumes of data. This Enterprise Data Strategy will continue to support DOE as we manage, store, secure, and leverage high-quality data, and ensure that America is positioned to lead the world in energy innovation and deployment. The impacts of our DOE Data Strategy will touch every corner of our department and the nation. In addition to leveraging powerful emerging technologies such as artificial intelligence and machine learning, this Data Strategy will further enable our data to fuel DOE's leadership in research and development and drive national priorities for decades to come.

Sincerely,

A handwritten signature in black ink that reads "David M. Turk".

**David M. Turk**  
Deputy Secretary  
U.S. Department of Energy

# Letter from The Chief Information Officer



The Department of Energy's robust mission space spans energy policy, nuclear security, classified and open science, energy delivery, and environmental management. Across every aspect of our work, innovation is key to the discoveries that will shape our energy future. Just as the varying security needs across DOE's mission space present unique cybersecurity challenges, the different Offices, Power Marketing Administrations, National Labs, Field Sites and Plants within DOE both produce and use data in ways that require special considerations to help accelerate innovation and achieve their mission.

This Enterprise Data Strategy reflects the commitment we've made in our DOE Office of the Chief Information Officer Information Technology Strategic Plan to strengthening data governance and promoting the use of data for strategic decision-making. I am pleased that representatives from so many different elements of the Department have come together to collaborate, share their perspectives, and build this strategy. I am confident that it will enhance data capture, processing, protection, privacy, confidentiality, handling, and analysis, maximizing the impact of DOE's rich data resources. At DOE, high quality, trustworthy data drives decision-making across all mission areas. Its impact is only amplified by the volume of data generated from increasingly complex physical systems and computer models enabled by high performance computing.

Our unique capabilities to leverage our vast data resources to benefit the American people underscore the importance of implementing this Enterprise Data Strategy, and I congratulate everyone who helped to make this happen.

Sincerely,

A handwritten signature in dark ink, appearing to read 'A. Dunkin', with a stylized flourish at the end.

**Ann Dunkin, P.E.**  
Chief Information Officer  
U.S. Department of Energy

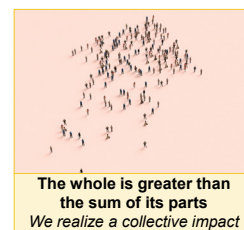


# Letter from The Chief Data Officer



To meet the needs of the Department of Energy's mission, data must serve as a strategic asset and a critical enabler for our enterprise. In the digital era, data is the lifeblood that unlocks the art of the possible, empowering us to achieve groundbreaking advancements in Artificial Intelligence (AI), analytics, and high-performance computing. Investing in our data foundations, including strong data management and governance frameworks, data architecture and platforms in the next 3 plus years will enable us to maximize the value of data for informed decision-making. I am excited to present DOE's Fiscal Year 2025-2028 Enterprise Data Strategy, which demonstrates how harnessing the full potential of our collective genius will drive measured actions, scaling impact across energy, science and innovation, safety and security, and mission operations.

My first-year focus as DOE's Chief Data Officer (CDO) has been on developing an Enterprise Data Strategy that reflects the whole of DOE, encompassing our extensive network of DOE Program and Staff Offices, Power Marketing Administrations, National Labs, Field Sites, and Plants. As CDO, I seek to orchestrate and convene best practices, identify opportunities to scale existing excellence, and catalyze collaboration. But it is our collective strength at DOE that drives our success. ***The whole is greater than the sum of its parts.*** Integration is key across DOE's complex functions, domains, and missions to realize our Data Strategy goals.



DOE's Enterprise Data Strategy will use targeted, priority use cases to demonstrate the value of implementing the Goals and Objectives with an emphasis on ***measuring what matters***. The core principle of this implementation approach is to ***not reinvent the wheel*** – a focus on executing activities and elevating what works, highlighting existing successes, and learning from past experiences. The initial phases of implementing the Strategy will involve aligning use cases with our strategic mission priorities to enable data-driven digital transformation. As we revisit and scale the incremental successes we achieve, we will continuously illustrate the value proposition. With this approach, DOE will showcase our progress and the impact of a shared data vision at every step.



Throughout the development of the Enterprise Data Strategy and internal Implementation Plan, we aimed to ensure representation across the enterprise. Our first step was to recharter the DOE Data Governance Board, codifying a robust representative body. Next, Integrated Working Groups built upon this collaborative spirit, and we extend our gratitude for the valuable contributions made through these data communities and forums. This Enterprise Data Strategy would not be possible without these partners across DOE – as we know, ***it takes a village***. We remain committed to fostering collaboration and welcome further engagement, recognizing that our collective strength lies in our shared efforts.

Moving forward, we are thrilled about the continued and emerging partnerships that are essential to address today's energy needs and pave the way for tomorrow's advancements. We encourage our partners and data leaders across the Department to engage in the robust collaboration networks we've built to execute our Enterprise Data Strategy and its tactical Implementation Plan. Together, we will fuel mission success across DOE, embodying our fifth core principle that ***all boats rise with the tide***.



Sincerely,

**Robert King**  
Chief Data Officer  
U.S. Department of Energy

U.S. Department of Energy

# Enterprise Data Strategy

## EXECUTIVE SUMMARY

The DOE Enterprise Data Strategy sets a comprehensive framework to collaboratively leverage data across the Department, aligning with our mission and future goals by prioritizing value-driven activities and strategic investments to enhance informed decision-making.

### • Vision •

Collaboratively unlock the full potential of DOE's complex data landscape to drive informed decisions that scale impact across energy, science & innovation, safety & security and operational missions.

## Guiding Principles



*Treat Data  
as a Strategic  
Asset*



*Align to  
Community Best  
Practices*



*Measure  
& Solve for  
What Matters*



*Embrace  
Incremental  
Progress*



*Foster  
Continuous  
Learning & Data  
Fluency*



*Design  
for Scale &  
Collaboration*

## Goals & Objectives

01



### ESTABLISH & SCALE FEDERATED DATA GOVERNANCE & STEWARDSHIP

- 1.1** Implement coordinated data governance structures across DOE to mature our governance model & better integrate data communities.
- 1.2** Define & promote frameworks that advance data management & quality practices at an enterprise scale.
- 1.3** Partner with DOE, Federal & industry thought leaders to elevate Department-wide data security controls to protect our Nation & critical assets.
- 1.4** Mature & regularly evaluate data stewardship practices.

02



### BUILD & MAINTAIN DATA ARCHITECTURE & INFRASTRUCTURE FOUNDATIONS

- 2.1** Promote shared data reference architectures, services & solutions that provide the scalability & flexibility required for advanced analytics & AI capabilities.
- 2.2** Adopt an API-first approach enabling system interoperability to secure & scale data access & discovery.
- 2.3** Accelerate processes for onboarding & deploying emerging capabilities that improve accessibility, usability & security of data, analytics & models.
- 2.4** Unify identity & access management solutions to enable collaboration through seamless governance & security controls.

03



### STRENGTHEN CAPABILITIES THAT IMPROVE DISCOVERABILITY, USABILITY, QUALITY & TRUSTWORTHINESS OF OUR DATA

- 3.1** Deploy & maintain an enterprise data catalog that serves as DOE's comprehensive data inventory.
- 3.2** Scale metadata standards to drive data consistency, discoverability & interoperability.
- 3.3** Promote best practices that accelerate access to data for authorized Department & public stakeholders.

04



### FOSTER & SUSTAIN OUR NEXT GENERATION DATA & ANALYTICS WORKFORCE

- 4.1** Define data skills, competencies & professional development pathways to mature DOE's data & analytics workforce.
- 4.2** Define standard roles & responsibilities to build a strong talent pipeline of critical data & analytics practitioners.
- 4.3** Enhance existing department-wide Communities of Practice to accelerate knowledge sharing & upskilling.
- 4.4** Strengthen institutional knowledge management, sharing & collaboration across data & analytics communities.

05



### ENSURE THE TRUSTED & SCALABLE USE OF DATA FOR AI

- 5.1** Strengthen practices, integration points & tools to promote discoverability for development & governance of AI.
- 5.2** Implement data management & governance practices for effective & appropriate use of data in analytics & AI.
- 5.3** Build transparent practices in end-to-end data & AI lifecycles to safely, securely harness AI.



# Introduction

The mission of the U.S. Department of Energy (DOE) is to address energy, environmental and nuclear challenges through transformative science and technology solutions to ensure America's security and prosperity. DOE's Program and Staff Offices, Power Marketing Administrations, National Labs, Field Sites and Plants - collectively referred to as DOE "components" in the remainder of this document - catalyze the transformative growth of emerging technology with basic and applied scientific research, as a cornerstone of U.S. economic prosperity. DOE hosts and supports a broad ecosystem of unique capabilities that advance our mission, producing or providing vast amounts of valuable data in many forms. Obtaining sound information from and safeguarding the many different types of data that DOE creates, hosts and supports is essential to effectively fulfilling our mission.

We stand at an inflection point for optimizing and managing data assets. Recent, rapid advancements in AI capabilities present exponential and exciting opportunities to unlock the power of data to advance mission priorities like never before, from revealing profound scientific insights to optimizing facility operations, mitigating security threats, and building the energy economy. These emerging capabilities have also exposed the various challenges of translating data into impactful results, including managing the data lifecycle, ensuring appropriate data use, providing effective security and enabling advanced analytics. As DOE explores new ways to take advantage of growing AI capabilities within the context of our complex data landscape, we must continue to protect data privacy, embed mechanisms to enable trusted and transparent AI practices, safeguard our strategic assets and build robust infrastructure that mitigates risk and continues to effectively steward investments.

DOE must remain a thought leader as we navigate this critical crossroad. Our role extends beyond being a thought leader to joining our Federal, academic, private and public partners in establishing best practices and provisioning the resources to mature the data foundations required to realize the power of emerging capabilities. Our leadership relies on strong strategies and coordinated governance structures that enable effective data and analytics infrastructure, foster innovative solutions to challenges and promote workforce readiness. With this culture, we can unlock the full potential of our data to enable new opportunities to inform decisions that impact the safety, security and scientific innovation shaping the future of our nation and world.

## Enterprise Data Management at DOE

The Department's scope and scale of managing data as a strategic asset spans defense and national security, research, development, power production, grants, oversight and innovation responsibilities. Our data is federated based on highly nuanced and often interconnected mission areas, data domains and organization-specific requirements. There are many considerations involved in determining how to most effectively govern the data DOE creates, manages, or hosts based on these complex tiers of requirements. For example, there are variations in how data is managed among our National Laboratories, Field Sites and Plants and across the broader DOE ecosystem; each data-generating effort must consider contract requirements, cross-government research agreements and mission needs.

In this document, any references to "data" and "our data" includes all data that DOE is accountable for stewarding, which includes DOE-generated data, data DOE is licensed or provided usage rights to, and DOE-supported publicly shared data.

### Enterprise

The term "enterprise," both in the EDM Program name and throughout this strategic document, embodies efforts impacting multiple or all DOE components.

### Component

Within the context of this Data Strategy, a "component" is a term that includes the entire ecosystem of DOE's Program and Staff Offices, Power Marketing Administrations, National Labs, Field Sites and Plants.



Realizing DOE's mission to its fullest potential will require us to embrace a mindset that respects the unique intricacies of managing business, scientific research or other data and while also embracing the need to explore and act on the interconnectedness of our various data domains. Formally defining the granular tiers of data domains and related governance processes for DOE-managed data is a significant undertaking and a key priority within this Data Strategy, serving as a foundational pre-requisite to enable many other strategic objectives. In the interim, the informal data domains noted in Figure 1 can be used to understand common categories of DOE "enterprise" data domains considered in scope for this Data Strategy.

The appointment of a Chief Data Officer (CDO) for DOE positions the Department to advance coordinated governance and security of its complex data landscape. The CDO is charged with coordinating across agency leaders to "ensure the successful execution of the agency's data management responsibilities."<sup>1</sup> The CDO established an Enterprise Data Management (EDM) Program to serve as one mechanism to deliver on DOE's mission. The EDM Program connects DOE components with the guidance, tools and support to respond to legislative requirements, like those defined in the *Foundations for Evidence-Based Policymaking Act of 2018 (Evidence Act)*,<sup>2</sup> and seamlessly integrate enterprise data goals and best practices into existing priorities and operations. The EDM Program supports the operations of the Department's Data Governance Board (DGB). The DGB is DOE's official, collective data governance body, required by the Evidence Act to "set and enforce priorities for managing data as a strategic asset to support the agency in meeting its mission."<sup>3</sup> The CDO serves as the Chair of the DGB, which includes senior data Federal employee representatives from across DOE support and program offices. The DGB advises on data-related matters across the Department and reports recommendations and decisions to Departmental leadership.

The EDM Program (see Figure 2) also operates to serve in a broader crosscutting role for DOE data-related priorities, working closely with the DGB to support the CDO to convene DOE components to design and execute strategic data management, governance and analytics objectives at an enterprise scale. The EDM Program is a vital connector for data-related communities across DOE, promoting the adoption of best practices and thought leadership to address complex challenges. A core principle of the program is to elevate and scale existing examples of excellence across DOE components to avoid "reinventing the wheel," accelerating problem solving and innovation.

<sup>1</sup> *About Chief Data Officers | Federal CDO Council*

<sup>2</sup> *The Foundations for Evidence-Based Policymaking Act (Evidence Act) of 2018 (2018).*

<sup>3</sup> *Office of Mgm't & Budget, Exec. Office of the President, OMB M-19-23, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance | (2019).*

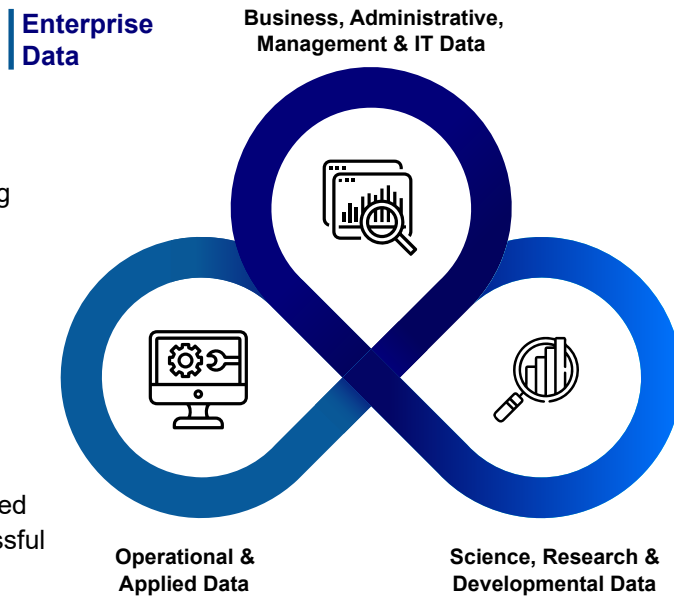


Figure 1 | Three core data domains span the robust mission areas of DOE components, all unified under the comprehensive term "Enterprise Data."

### EDM Program Operating Pillars

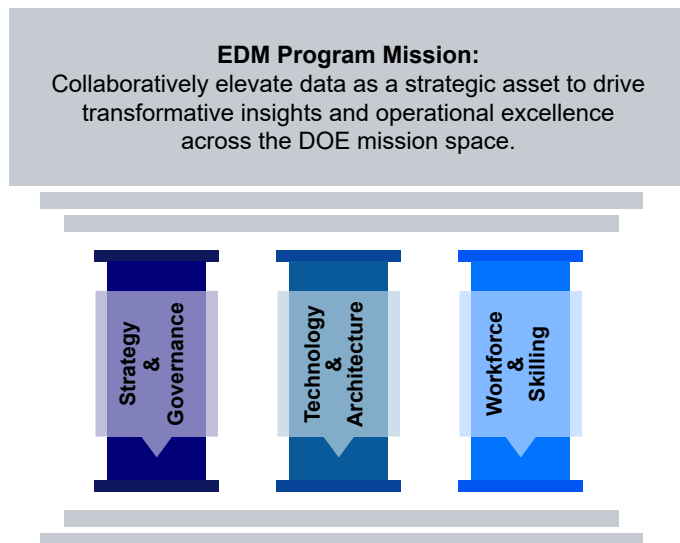


Figure 2 | EDM Program mission statement and operating pillars reflect the aspirations and initiatives that will remain central to supporting DOE to realize the success of goals outlined within this Data Strategy.

## Our Collective Approach to Creating the FY25-28 DOE Enterprise Data Strategy

DOE's mission and culture is rooted in creating transformational knowledge, insights and solutions across a wide range of scientific, technical and professional domains. To advance our mission, we steward 17 National Laboratories, host world-class scientific facilities and world-leading computing capabilities, are the nation's largest supporter of basic research in the physical sciences, and build collaborations with Federal, industrial, academic and international partners. Our commitment to remaining at the forefront of scientific research, national security, energy and environmental cleanup efforts drives us to construct a comprehensive data strategy for the entire enterprise.

This Data Strategy goes beyond compliance to be a unifying vision and roadmap for the needs and aspirations of our complex ecosystem, DOE components and stakeholder communities. Elements of our Data Strategy directly align to requirements and best practices defined in legislative statutes, national security requirements and administrative orders including but not limited to the *Foundations for Evidence-Based Policymaking Act (Evidence Act) of 2018*<sup>4,5</sup> and the *Federal Data Strategy*.<sup>6</sup> In addition to these requirements, developing a Data Strategy that resonates at an enterprise scale for an organization as nuanced and varied as DOE required intentional, extensive inputs from data, analytics and AI practitioners across every type of DOE component. **This document, and the Department's confidence in it, would not be possible without the invaluable contributions of hundreds of individuals across DOE's Program and Staff Offices, Power Marketing Administrations, National Laboratories, Field Sites and Plants and the full commitment of the Office of the Chief Information Officer (OCIO) and broader Departmental leadership.**

The dedicated engagement and candid conversations with these components across interviews, capability assessment surveys, workshops and other extensive discovery activities over a 6-month period were instrumental to the Data Strategy. Together, we generated the strategic goals, objectives and measures of success described in this document. Our goals and objectives aim to create tremendous value, support data-driven decision-making, create operational efficiencies and serve as baseline for meaningful investments that accelerate respective areas of mission success.

This Data Strategy sets a strong enterprise "North Star" to guide DOE components through enterprise-wide activities and investments. **Most importantly, our Data Strategy focuses on providing flexibility in subsequent implementation activities for DOE components to respond to the unique needs of respective mission areas and adapt to the inevitable unknowns to come with new innovations and external and environmental influences.** The corresponding implementation approach for this Data Strategy must accommodate the variations within DOE's data landscape. These considerations are detailed in the following section.

<sup>4</sup> *Office of Mgmt & Budget, Exec. Office of the President, OMB M-19-23, Phase 1 Implementation of the Foundations for Evidence-Based Policymaking Act of 2018: Learning Agendas, Personnel, and Planning Guidance | M-19-23 (2019).*

<sup>5</sup> *Related: OSTP, Exec. Office of the President, Ensuring Free, Immediate, and Equitable Access to Federally Funded Research (Aug. 25, 2022).*

<sup>6</sup> *Office of Mgmt & Budget, Exec. Office of the President, OMB M-19-18, Federal Data Strategy - A Framework for Consistency (2019).*





## Our Path Forward: Implementing the Enterprise Data Strategy

To navigate the intricacies of DOE's data and drive scalable progress towards our identified goals and objectives, our implementation activities must align to and account for the federated nature of the Department. Our implementation approach must continue to engage all stakeholders and account for factors such as but not limited to relationships between ever-growing and changing **data domains**; considerations of needs for **classified and unclassified environments**; ensuring consistency with **national security requirements**; and the balance of **promoting innovation with risk mitigation**.

DOE's implementation approach began in tandem with the development and publication of this Data Strategy. Extensive feedback from our months-long engagement with DOE components served as inputs to detailed activities, data types impacted, measures of success and other indicators that will guide progress towards objectives. DOE processes and forums like the enterprise DGB are being leveraged to regularly revisit progress against the Data Strategy and adapt our tactical approach and priorities as required.

DOE's CDO, Chief Information Officer, Chief Artificial Intelligence Officer, Chief Information Security Officer and other data leaders across all DOE Components, who engaged through the DGB, will continue to closely collaborate to define and sponsor Data Strategy implementation activities.<sup>7,8</sup> The EDM Program will be leveraged as a convening mechanism for data communities and will support the power of integrated, tactical teams of talented data stewards, data architects and other data and AI practitioners across DOE to develop and execute initiatives. Relevant DOE components will be engaged to collaborate on shaping, funding and executing strategic efforts. These initiatives will advance our strategic objectives and foster a more efficient and aligned data ecosystem.

This continued level of cross-DOE partnership in the implementation phase of our Data Strategy is essential given the interconnectedness of DOE managed data that supports mission activities. "A rising tide lifts all boats" serves to reflect this mindset that collectively fulfilling the priorities in this strategy to mature data foundations at both an individual DOE component and enterprise scale will create cascading impacts that benefit all mission areas.



<sup>7</sup> [Additional Information on CDO, CIO & CISO | Office of the Chief Information Officer.](#)

<sup>8</sup> [CAIO Office of Critical and Emerging Technologies.](#)



## Goal 01

# Establish & Scale Federated Data Governance & Stewardship

Goal 1 is focused on establishing coordinated data governance operating models and data management practices across all tiers of the Department. These essential elements will enable DOE to better align individual data initiatives to create a federated, interconnected enterprise data ecosystem that effectively drives cross-cutting mission and operational priorities. Coordinating intentional forums for leaders and practitioners across our data, analytics and AI communities is necessary for consistent knowledge sharing and elevating partnership opportunities for common needs. Initiatives will draw from successful and proven federal, industry and cross-sector data community expertise for data governance, management and security practices. Actively partnering across all DOE components is key to designing and implementing enterprise policies, frameworks and standards. It allows DOE to balance the need to tailor best practices to unique DOE component data ecosystems and operating constraints while remaining flexible and agile enough to adapt to constantly evolving internal and external environmental factors.

### Data Governance

Data governance is defined within the context of this Data Strategy as an accountability framework which enables data as a strategic asset to meet stakeholder needs related to the valuation, creation, consumption and integrity of data. It establishes data management roles and responsibilities that define how decisions are to be made about data and support data resource allocation.

### Objective 1.1

Implement coordinated data governance structures across DOE to mature our governance model & better integrate data communities.

We will design a federated data governance operating model that outlines clear roles and responsibilities at each tier of the enterprise, spanning governance boards to individual data-related practitioners like data stewards. We will engage DOE components, working across domains and communities to identify, define and integrate governance operations where they exist today into the enterprise model. The EDM Program will support efforts to adapt and implement operating model frameworks where gaps may exist in individual DOE components. Scaling data governance structures to the DOE enterprise ensures the required flexibility to accommodate component and domain specific realities, and balances reducing risk while driving innovation and decision-making. Creating and maturing real-time engagement and communication channels, both formal and informal, is key to lowering barriers to knowledge sharing, decision-making and innovation across our data communities (see *Objective 4.4*). Objective 1.1 aims to consistently engage with relevant stakeholders in appropriate forums to better identify common needs, build on success and ideas of other DOE components to accelerate value and reduce data-related risks.

### Objective 1.2

Define & promote frameworks that advance data management & quality practices at an enterprise scale.

We will establish and maintain data management frameworks, policies and standards that elevate best practices from within DOE components, industry and cross-sector data communities to promote interoperability and flexibility across the range of DOE data and analytics systems and data types. As in Objective 1.1, we will design frameworks to accommodate unique implementation considerations to ensure enterprise scalability. This includes varied data classification levels or domain-specific handling and access control requirements. Scope also includes aligning retention and disposition controls adoption for DOE data to the data lifecycle to enhance the understanding and availability of appropriately maintained records that are relevant, usable and high quality to support our missions. We will assess and prioritize efforts to improve data quality and management practices, including the consideration of data ownership and stewardship, usage rights, provenance and relevant data and metadata standards. Objective 1.2 accelerates implementation of data management best practices, increases reusability and fit-for-purpose use of DOE datasets and improves the quality and trustworthiness of our data.

### Objective 1.3

Partner with DOE, federal & industry thought leaders to elevate Department-wide security controls for data to protect our nation & critical assets.

We will collaboratively partner with cybersecurity experts and thought leaders across DOE, federal councils, industry and academia to continuously elevate, monitor and adapt NIST-defined security controls to protect DOE's and our nation's critical, sensitive data assets. Our efforts must enable legislative, administrative and mission-driven data sharing requirements while mitigating ever-growing security risks. We will proactively engage DOE data and cybersecurity leaders and broader communities in relevant forums to elevate and promote strong data protection and security practices, leveraging frameworks like Zero Trust. Dedicated, cross-functional teams of data and cybersecurity practitioners are essential to establishing the required practices and solutions for continuous identification and mitigation of evolving security threats. This objective decreases data-related risks to enhance DOE-wide preparedness, ensuring readiness for adverse events and limiting disruption to mission outcomes for DOE, our partners and the public.

### Objective 1.4

Mature & regularly evaluate data stewardship practices.

We will build on *Objective 1.1*, expanding robust processes, fostering communities and defining formal responsibilities for DOE data stewards. Data stewards play a critical role in strengthening data management and security outcomes and ensuring compliance with applicable data-related policies. Regular evaluation and evolution of stewardship frameworks at both the enterprise level and for specific domains will benefit from cross-pollination and collaboration among DOE components. Effective practices that benefit one domain, like initiatives that generate reusable assets, skilling programs and integrated working groups that provide tangible value and practical guidance, can be more readily shared and tailored for other applications. Engagement channels will remain transparent, with continuous feedback and monitoring mechanisms to evaluate both individual satisfaction and enterprise impact to data-related key performance indicators. Objective 1.4 is important because successful implementation of the Data Strategy objectives relies on a foundation of effective stewardship practices to achieve data quality, discoverability and security outcomes vital to downstream data use for advanced insights and decisions.





## Goal 02

# Build & Maintain Data Architecture & Infrastructure Foundations

Goal 2 prioritizes policies, systems and holistic solutions that will enhance the efficiency, flexibility, security and reusability of DOE supported data environments. We are committed to continuously evaluating opportunities to increase the interoperability, consistency, quality and security of the DOE data ecosystem. To do this, we will ensure that DOE supported data platforms and infrastructure can scale and respond to the rapidly evolving technology and use case landscape while promoting effective integration, interoperability and reusability of existing resources. Efforts will establish enterprise policies that drive best architecture practices, maximize shared service offerings for common data and analytics use cases, and ensure we have the right technology foundations in place to maintain DOE's ability to lead innovation at the pace our global challenges require. Objectives aim to foster a next-generation technical environment that better enables seamless collaboration between DOE components, maximizes the allocation of resources to mission activities and optimizes the mechanisms for teams to onboard and access emerging capabilities.

### Objective 2.1

**Promote shared data reference architectures, services & solutions that provide the scalability & flexibility required for advanced analytics & AI capabilities.**

We will promote the use of data architectures, at enterprise-wide to component-specific scales, that ensure DOE supported data and analytics systems are appropriately standardized, interoperable and scalable to advance mission needs. Solutions and assets will prioritize best-practice technology principles and standards, like machine-readability, rather than prescribing specific tools. This will practically accommodate unique implementation considerations and promote the core capabilities required to enhance collaboration and efficiency in data utilization and management. Shared, enterprise solutions will promote modern best practices in data analytics and data curation, strengthened with privacy-preserving technologies and automated data flow and management. Agile approaches and maximum transparency will be encouraged for any DOE components developing and maintaining DOE infrastructure for data and analytics. Architectures will seek to integrate as appropriate with comprehensive suites of data science and analytics tools to provision capabilities that increase the availability and usability of data for advanced, mission insights and decisions. Objective 2.1 will drive the standardization and promotion of the best architecture practices needed to improve data reusability and system interoperability that advance DOE's mission needs.

### Objective 2.2

**Adopt an API-first approach enabling system interoperability to secure & scale data access & discovery.**

We will prioritize Application Programming Interface (API) first approaches and the common practices and standards required to deploy them for enterprise code and development projects. Promoting interoperable systems enables automation, resiliency and reusability that can expand integration and data sharing capabilities at-scale across DOE systems. We will engage cybersecurity teams and leaders to ensure necessary authorization and access management protocols as critical elements of workflow enablement. Our approaches will enable effective governance practices to enhance downstream quality. Objective 2.2 reflects DOE's commitment to fostering efficient data exchange and security while enhancing the discoverability of our data. APIs and other effective technical solutions will reduce the resources and time required to deploy desired data and analytics capabilities and improve data quality, usability, security and sharing of data.

### Objective 2.3

**Accelerate processes for onboarding & deploying emerging capabilities that improve accessibility, usability & security of data, analytics & models.**

We will promote optimization of pathways for our teams to rapidly explore, assess and implement effective innovative and emerging data-related solutions into DOE environments. This objective aims to ensure DOE's processes allow the agility required to take advantage of advances in the rapidly evolving technology landscape. We will facilitate collective efforts across DOE components with common priorities and limitations throughout the deployment lifecycle of data and advanced analytics solutions. These may span from procurement through Authority to Operate approvals, user training and onboarding and self-guided resources that support data or product interface usage questions and issue resolution. DOE will promote the periodic review and optimization of modern computing platforms to continuously identify enterprise opportunities to collaboratively reduce costs, improve performance and effectively use high value datasets. To support enterprise scalability, we will evaluate opportunities to pilot and deploy semantic tools like ontologies, knowledge graphs and large language models. Objective 2.3 promotes streamlined processes and supportive self-service resources to ensure teams can keep pace with best-in-class capabilities that advance mission outcomes while maintaining essential security and compliance processes.

### Objective 2.4

**Unify identity & access management solutions to enable collaboration through seamless governance & security controls.**

We will evaluate opportunities for streamlining authentication and authorization solutions to reduce barriers for collaboration between DOE components and with external partners. Initiatives include exploring a department-wide collaboration platform to provide identity and access management solutions as well as evaluating ways to improve seamless yet controlled sharing of complex and sensitive datasets, including but not limited to protective solutions like multi-factor authentication. Objective 2.4 is important not only because of *Executive Order 14028*<sup>9</sup> and the subsequent centralization of identity management systems directed in *OMB Memo M-22-09*,<sup>10</sup> but because it is essential for increasing the value of common platforms. Unified identity and access solutions are critical for federating enterprise infrastructure and governance to enable secure, scalable and trusted collaboration mechanisms for classified and unclassified use cases.

<sup>9</sup> Exec. Order No. 14028, 86 Fed. Reg. 26633 (May 12, 2021).

<sup>10</sup> Office of Mgmt & Budget, Exec. Office of the President, OMB M-22-09, *Moving the U.S. Government Toward Zero Trust Cybersecurity Principles* | M-22-09 (2022).



## Goal 03

# Strengthen Capabilities that Improve Discoverability, Usability, Quality & Trustworthiness of Our Data

Goal 3 encompasses DOE's continued commitment to making data findable, accessible, interoperable and reusable (FAIR),<sup>11</sup> as appropriate, for the Department, our partners and the public. Initiatives under this goal will improve DOE data discoverability and usage through effective metadata standards, best practice data management frameworks and user-friendly tools. Our aim is to enhance data sharing pipelines to provide more timely promotion of data to relevant internal and public-facing environments to better support mission progress and societal impact. To succeed, we will coordinate the complex alignment of metadata and quality standards across the enterprise and within our various data domains, ensuring we effectively maintain the integrity of our data sources. This enables us to accelerate innovation and advance a culture of "share first" and transparency.

### Objective 3.1

Deploy & maintain an enterprise data catalog that serves as DOE's comprehensive data inventory.

We will collaborate with DOE components to design and implement an enterprise data catalog that will enhance data findability, accessibility and sharing capabilities. The enterprise data catalog will maintain DOE's comprehensive data inventory to enable better management and protection of data assets. The catalog will enable compliance with various Federal legislative statutes, executive orders and guidance like the *Evidence Act*, *Geospatial Data Act of 2018*,<sup>12</sup> *EO 13859*<sup>13</sup> and *M-13-13*.<sup>14</sup> The aim is to develop a flexible, scalable data catalog implementation that enables tailored connections to individual agency components to leverage existing investments in data management. A successful catalog will interlock with *Objective 1.4* to map the roles of domain data stewards to responsibilities in metadata management activities. We will promote features that reduce manual burden for data stewards and automate processes to maintain the recency and quality of metadata. Objective 3.1 actively considers interdependent objectives to establish standards, leverage APIs for interoperability and provide clarity of data roles and responsibilities to deploy a data catalog that drives value and meets requirements at an enterprise scale. This capability will directly increase the connectedness of our data communities, and the overall quality and trustworthiness of DOE data.

### Objective 3.2

Scale metadata standards to drive data consistency, discoverability & interoperability.

We will collaborate with domain data stewards across DOE components to establish and promote metadata standards that accelerate data classification, discoverability and interoperability to strengthen informed decision-making across DOE mission areas. Efforts to design common metadata standards will account for nuance between enterprise-level versus domain-level requirements and build on existing best practices and principles from Federal and legislative requirements (e.g., Zero Trust Architecture, *Open Government Act*<sup>15</sup>), DOE policy, industry and academia. We will promote standards that enable interoperability and compliance while providing flexibility and autonomy for data communities to tailor standards to domain-relevant requirements. Objective 3.2 will strengthen foundational metadata standards that serve as pre-requisites for achieving broader Data Strategy objectives. Strong metadata management is essential for achieving reduced time to insights and enhanced data quality and access outcomes at scale.

### Objective 3.3

Promote best practices that accelerate access to data for authorized Department & public stakeholders.

We will assess and promote best practices to our partners and stakeholders that enhance the quality, robustness, timeliness and security of shared data for both DOE mission and public use. We will communicate with internal and external stakeholders to share opportunities and frameworks along the end-to-end data and analytics pipeline that improve data sharing while safeguarding assets. Materials may focus on promoting mechanisms that improve principles such as: machine readability of metadata; usage of persistent identifiers; data tagging frameworks for classification and anti-corruption; and standardization of protocols to access and retrieve desired metadata fields. We will evaluate DOE-supported open data platforms and integrations with Federal repositories (e.g., Data.gov, GeoPlatform.gov) to identify opportunities to streamline and automate data classification, tagging, securing and promotion activities. Frameworks will embed community-centered design principles to encourage maximizing the usability of published data in accessible, intuitive interfaces and ensure data shared for public consumption has clear value and demand to best steward DOE resources. Materials published through Objective 3.3 will promote best practice frameworks that will aid in optimizing data sharing processes and investment efficiencies to ultimately strengthen collaboration and mission outcomes.

<sup>11</sup> *FAIR Principles - GO FAIR (go-fair.org) (2024)*

<sup>12</sup> *Geospatial Data Act of 2018, 43 U.S. Code Sections 2801-2811 (2018).*

<sup>13</sup> *Exec. Order No. 13859, 84 Fed. Reg. 3967 (Feb. 11, 2019).*

<sup>14</sup> *Office of Mgm't & Budget, Exec. Office of the President, OMB M-13-13, Making Open and Machine Readable the New Default for Government Information | M-13-13 (2013).*

<sup>15</sup> *Open Government Act of 2007, 5 U.S. Code Section 552 (2007).*





## Goal 04

# Foster & Sustain Our Next Generation Data & Analytics Workforce

Goal 4 acknowledges that concerted investments in creating and maintaining a strong data and analytics workforce are essential to achieving all other strategic objectives and overall mission success. Initiatives will strengthen elements throughout the career experience for data-related practitioners, from initial recruitment and hiring to continuous development and upskilling offerings. We will also focus on driving a collaborative culture through purposeful revitalization of grassroots and formal data community operations to accelerate knowledge sharing and innovation. Objectives in this goal will improve the resiliency and power of our workforce to leverage data through analytics, including AI, and position DOE to continue its leadership in solving some of the most pressing national and global challenges of our time.

### Objective 4.1

**Define data skills, competencies & professional development pathways to mature DOE's data & analytics workforce.**

We will design holistic career pathways with elements like fit-for-purpose trainings and professional development programs, including clear incentives to drive strong recruitment and long-term retention of talented data and analytics practitioners. Career pathways will be anchored in DOE and industry-standard, baselined skills, with mechanisms to measure tangible progress and value to program participants. Integrated DOE groups will work to identify common skilling needs across components and build partnerships to facilitate scaled efficiencies for investments like licensing and training platform access. Training or broader development initiatives will connect prioritized growth areas to direct application in current roles or new opportunities. Objective 4.1 aims to develop rewarding, mission-centered career journeys that support individual growth in data fluency, continuing to position DOE as industry leaders in this dynamic data and analytics environment. These efforts will enhance job satisfaction and retention, bolster priority skills and strengthen the talent pipeline from recruitment to conversion and retention, fostering a culture of confident practitioners that understand the high mission impact they drive in their roles.

### Objective 4.2

**Define standard roles & responsibilities to build a strong talent pipeline of critical data & analytics practitioners.**

We will convene cross-functional data, analytics and human capital teams to standardize common roles and responsibilities in modern position descriptions, using these as mechanisms to strengthen the competitiveness of DOE roles and overall talent pipelines. Streamlining position descriptions complements *Objectives 1.1* and *1.4*, as formalizing essential data management, stewardship, architecture, engineering, librarian, scientist and other roles will enable the proper recognition of positions that take on these responsibilities. Efforts will account for the specialized nature of DOE roles and consider tailored approaches where these responsibilities may need to be integrated as elements of a broader role to appropriately staff highly desired skills. These exercises will mutually benefit talent managers to increase their data fluency to support recruiting foundational and emerging data and analytics talent. Revitalizing standard roles and responsibilities through Objective 4.2 will allow DOE components to effectively fulfill workforce needs, design support resources and growth opportunities (see *Objective 4.1*) and assess and reward performance for employees.

### Objective 4.3

**Enhance existing department-wide Communities of Practice to accelerate knowledge sharing & upskilling.**

We will revitalize DOE's Community of Practice (CoP) models to maximize knowledge sharing and collaboration across DOE components. These efforts aim to build a comprehensive, discoverable inventory of data- and analytics-related CoPs, working groups and other communities of interest. We will also evaluate opportunities to combine initiatives with similar goals. We will provide guidance for CoP operating model best practices that clarifies operational expectations and recommends ways to balance open participation with targeted engagement to effectively realize their desired impact. Success of CoPs is dependent on formal and informal communication channels (see *Objectives 1.1* and *4.4*) to effectively exchange ideas, network, host shared collaboration spaces and celebrate success stories that can be reused at scale. Objective 4.3 elevates the excellence and innovation already at DOE through communities of practitioners closest to our mission and operations. CoPs scale our ability to solve problems, connect our workforce and lower barriers to informal upskilling opportunities provided through participation in these communities.

### Objective 4.4

**Strengthen institutional knowledge management, sharing & collaboration across data & analytics communities.**

We will leverage policy, processes and intelligent solutions to strengthen knowledge management for collaboration across DOE. Policies will promote best practices, such as maintaining system documentation and data dictionaries or catalogs, to increase data fluency and self-service knowledge exchange capabilities. We will explore advanced intelligent solutions for codifying institutional knowledge and automating manual processes to reduce risk of knowledge loss and accelerate time to insights. Embracing citizen development philosophies, we aim to enable practitioners at all skill levels with tools and methodologies that reduce barriers accessing and utilizing data required to effectively execute their roles. Objective 4.4 initiatives are closely tied to other strategic objectives given the dependencies of strong knowledge management processes and tools to elements like data management standards; unified identity management; accessible shared environments; and cross-DOE forums that actively work to deliver value and outcomes. Objective 4.4 seeks to establish best practices and solutions that will encourage a culture of maximal appropriate data sharing, remove barriers to collaboration and improve the resiliency of DOE's workforce to collectively sustain our ability to achieve critical mission outcomes.



## Goal 05

# Ensure the Trusted & Scalable Use of Data for AI

Goal 5 invests in the necessary foundations to provide curated, accessible and quality data pipelines that enable the development and application of AI tools to provide insight and advance the DOE mission. These pipelines directly fuel the ongoing advancement of robust and trustworthy AI for scientific discovery, resilient energy futures, and ensuring the safety and security of our Nation. As a recognized leader in science and innovation, DOE will continue to lead and partner with Federal, academic and industry stakeholders to conduct red teaming for system security, design evaluation frameworks, drive research in AI for energy and global challenges, and lead workforce development initiatives. We must also mitigate the risk of potential misuse and threats these technologies potentially perpetuate. Success in these efforts requires fostering a culture that recognizes the end-to-end data and AI lifecycle as interdependent and holistic practices. Promoting best practices in data capture, processing, protection, privacy, confidentiality, handling and analysis is essential for developing AI models that drive critical decisions and scientific innovations. AI also holds the potential to accelerate and enhance the efficiencies of foundational data management activities for organizations. The objectives outlined in Goal 5 are focused on initiatives at the intersection of DOE's data and AI priorities that strengthen our cross-functional culture and operations.

### Objective 5.1

**Strengthen practices, integration points & tools to promote discoverability for development & governance of AI.**

We will collaborate with the Chief AI Officer (CAIO) and DOE components to design systems with intuitive processes and resources, equipping practitioners to effectively navigate and use our complementary data and AI inventories. Enhancing the integration of data and AI inventories, tools and models into a more unified management system enables us to strengthen scientific, information technology and business processes, improving the discoverability of data and AI systems. A joint approach enables DOE to have coordinated platforms to manage these connected assets. Objective 5.1 aims to integrate design through implementation and maintenance activities to deploy and scale our shared data and AI inventories. This objective envisions simplified processes and advanced tools for DOE and specific components to improve asset management, enable crosscutting collaborations and the enhance our ability to effectively leverage our collective resources.

### Objective 5.2

**Implement data management & governance practices for effective & appropriate use of data in analytics & AI.**

We will implement initiatives that identify, enhance, adopt and scale processes or approaches to ensure data is efficiently collected, organized and maintained to reduce time to provision data for AI development. These efforts will include assessing opportunities to leverage analytics, including AI, in our data management and stewardship frameworks like automation of classification and sensitivity tagging recommendations for applications to data and records management and other methods to optimize practices that increase the reliability and usability of our data assets. Objective 5.2 efforts aim to accelerate data provisioning; scale data management systems to accommodate increasing use; remove manual or tedious data management activities for data stewards and other practitioners; mitigate security risks; and improve data quality to realize more effective and insightful decision-making processes.

### Objective 5.3

**Build transparent practices in end-to-end data & AI lifecycles to safely, securely harness AI.**

We will prioritize traceability throughout the data and AI lifecycle, from data acquisition to AI deployment, reducing the likelihood of biases and malicious actors exploiting our data and AI systems. Integrating accountability and transparency measures into existing infrastructure and designing frameworks or assets where gaps may exist for best practices in data quality and management will shape and scale criteria for the maturity, deployment and use of AI systems. We will execute these efforts in close partnership with the CAIO and DOE components and leverage existing Federal, industry, and cross-sector data community best practices, standards and requirements. Objective 5.3 initiatives will seek to result in data and AI systems that, throughout their lifecycle, mitigate the risk of bias; safeguard assets; ensure transparency and accountability for decisions; and foster increased trust among practitioners developing the systems, stakeholders impacted by outcomes, and the public.





U.S. Department of  
**ENERGY**™