



From Chaos to Clarity: Actionable Insights for Supporting Data Stewards as You Mature Data Governance

Kimberly Maestas, Co-Director of the Mission Data Stewardship Alliance at Los Alamos National Laboratory

10/24/24

LA-UR-24-31143



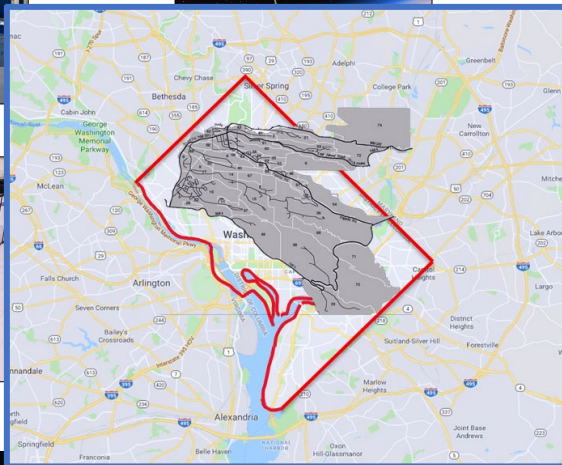
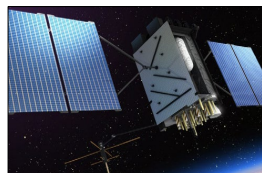
Today's Takeaways

BLUF By engaging in some key activities, we can support our data stewards today and lay the groundwork for future data governance initiatives.

Attendees will learn about

1. a new data stewardship effort at LANL and the role of a “data ambassador”
2. Predictable reasons data governance efforts fail and why working closely with data stewards sets up data governance to thrive
3. How to prepare for the next “Now What... Opportunity”

Los Alamos National Laboratory is a complex, dynamic system of people, facilities, and services



BY-THE-NUMBERS

40 square miles	Almost 900 buildings, 13 nuclear facilities
> 14,000 workers	1,800 students, 460 postdocs

Last updated FY22

DOE/NNSA

Triad Board of Directors

Laboratory Director's Office

Laboratory Director
Thomas Mason

Deputy Directorate,
Science, Technology, &
Engineering

Deputy Directorate,
Weapons

Deputy Directorate,
Operations

Intro to the Mission Data Stewardship (Midas) Alliance

- **Established May 2024**
- **Mission:** Support data stewards so they are positioned to leverage opportunities and lower barriers to deriving value from data.
- **Goals:** Establish a network of data stewards and build relevant models for data stewardship across the Laboratory.
- **Key Principles:** Data is an enterprise asset; Engage all stakeholders
- **Staff:** 2 full time co-directors
- **Funding:** Advanced Simulation and Computing (ASC) Program



Mike Ham
Mikeh@lanl.gov



Kimberly Maestas
kgotches@lanl.gov



Midas strives to unite LANL Data Efforts

- **Steward Advocacy:** Gather feedback from data stewards and funnel that feedback through data governance and other management channels. (“Data Ambassadors”)
- **Data Literacy:** Facilitate workshops and share resources to build long-term data stewardship muscle
- **Partnerships:** Collaborate with business teams and data delivery teams to streamline data management processes.



Dalle: King Midas building 'data stewardship muscle.'



The knockout blows for data governance can all be addressed with a strong foundation in data stewardship.



- “We had good intentions, but it turns out what we were doing wasn’t relevant to the businesses we were trying to help.”
- “We failed to quantify the impact of our efforts.”
- “There was a lack of resource commitment.”
- “We had poor incentives to motivate people to work with us.”
- “We had no authority or accountability.”

Fast forward 6 months... Imagine you have achieved exactly what you've been striving for.

Now What... Opportunity

We need to be able to say to our sponsors/leaders:

- “If this project can get this data, they can deliver this [].”
- “If we don’t remove this barrier to getting this data, this project will NOT deliver this [].”
- And this is what we need from you.



Photo Credit: PickPik.com



What are some worthwhile early focus areas for a data stewardship program?

(1) Find existing communities and projects, and make yourself useful

(2) If you think you are articulating strong value statements, think again

(3) Start building a culture to embrace changes in data management

Doing these three things now will not be wasted work.



Pilot: In Situ Data Stewardship in R&D

PROBLEM

- The practice of scientific research in the digital age has led to a significant increase in the amount of data generated and collected.
- Researchers are struggling to keep up with their program/funder demands and the growing demands of data management.
- Research program managers are similarly overwhelmed and unsure of what actions to take.
- Existing attempts to assign information professionals to research teams in the role of data concierges have not yet solved the problem.

SOLUTION

- Identify and support existing data stewards within research programs and projects.
- Formalizing the roles/responsibilities of data stewards within the teams, ensures data management strategies are implemented at the time of data creation, which leads to thorough documentation and data preservation for decades in the future.

APPROACH

- Conduct a series of workshops with the diagnostic, design, and simulation teams to understand their workflows, assess the maturity of their existing data stewardship practices, and formulate recommendations for in-situ stewardship.
- Working with the program on an action plan, including establishing requirements for reports and data management practices.
- We will consult with the program on an ongoing basis.

“ If someone needs to re-run your analysis in 5 years, can they? ”

If you think you are articulating strong value statements, think again



Dalle: Misaligned value proposition

- The process of implementing formal data stewardship will have far-reaching effects on an organization.
- In some cases, people will be asked to change how they work with data and how they interact with each other on activities involving data.
- Change takes place when people behave differently **because they recognize the value in doing so.**
- Therefore, if we want to help lead these changes, we need to get **REALLY** good at communicating value propositions.



Value Statement Test: Let's choose the best example

Formula: Organizations/Teams that do (X) *demonstrate business value improvements (Y).*

Test 1: Who cares the most about Y? (Must be the business)

Test 2: Can you describe how Y is a possible result of X?



Example 1: Organizations that formalize data stewardship responsibilities into daily routines *quickly view data governance as non-threatening rather than over and above the existing work effort.*

(Depends)

Example 2: Organizations that enforce stricter data management policies *assure themselves of lower levels of risk in data compliance assessments.*



Example 3: Teams that manage their data well *achieve higher-quality research, increased visibility, and the consequent benefits of enhanced citation rates.*

Fast forward 6 months again... Now what?

- develop awareness of the role of data stewardship and data stewards in different environments (research, business)
- Streamline data stewardship processes to support downstream analytics and AI use of data
- provide more forums for those working as data stewards to share knowledge and best practices through peer support and cross-disciplinary collaboration

Now What... Opportunity



Photo Credit: PickPik.com

“Great things are done by a series of small things brought together.” – Vincent Van Gogh



Contact Info

- Reach Kimberly at kgotches@lanl.gov
- Reach the LANL Weapons Program Data Governance Council/Mission Data Stewardship Alliance at WP-DGCRequests@lanl.gov

Additional Resources

Look for the Midas 2024 D3 poster – (U) Preparing LANL's Data for National Security AI applications: The Ambitious Vision of the Mission Data Stewardship (Midas) Alliance

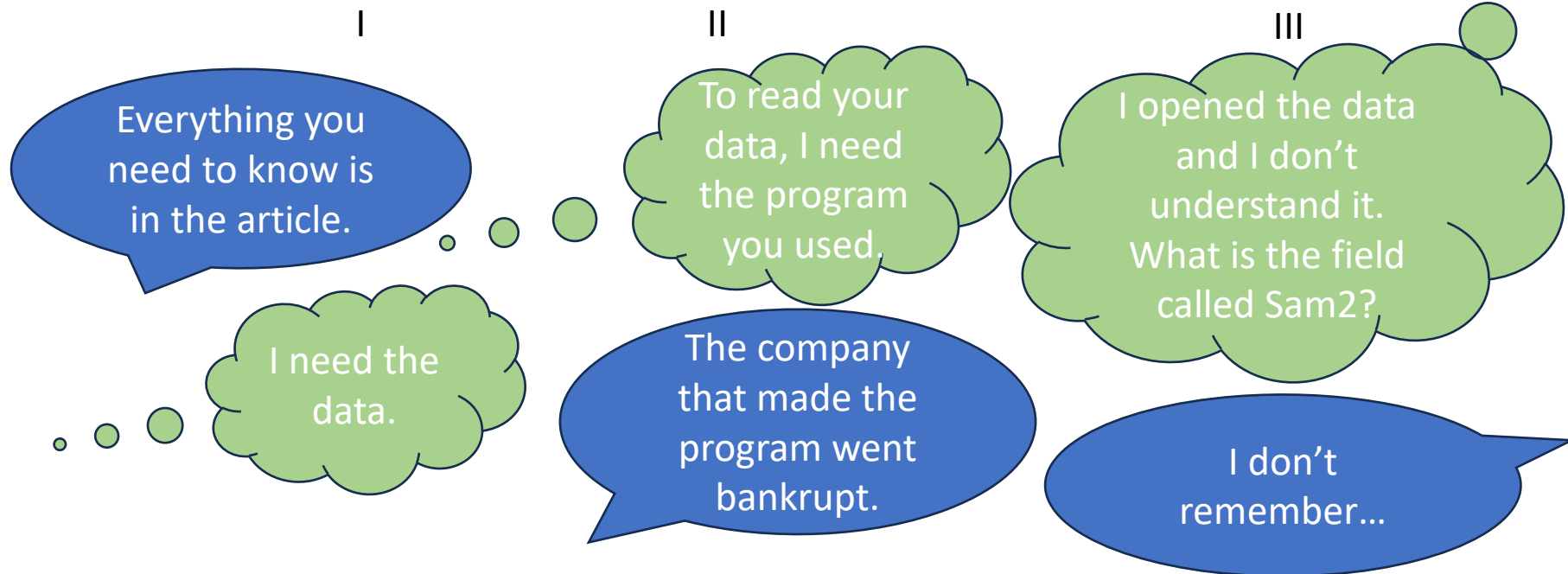
The stewardship model for research data should not be the same as the model for enterprise (business) data.

Derived from "Research vs enterprise: the data showdown" by San Cannon.

Research Data Activities	Enterprise (Business) Data Activities
Use data to better understand how the world works/ understand a concept.	Use data to better understand enterprise functions and improve business processes.
Data = an instrument of research and new knowledge	Data = \$
Questions, once answered, may never be asked the same way again.	Questions are usually repeated; Regular monitoring of processes are often crucial to business success.
Data is most often unstructured	Data is most often structured
Research data needs to be collected, captured, created, or contracted from the environment being studied. E.g. Raw data from an x-ray machine	Exists in applications or internal repositories
Project-based. Typically, each project has a unique data management plan	Function-based
Aim for consumers to understand the data used for the research to allow for potential replication or extension of the work.	Aim for consumers to understand context around the data so it may be interpreted and used correctly.
Approaches to manage the data are often project-dependent	There are well established "best practices" and "maturity levels" to describe how enterprise data are and should be managed

Aiming to create a data culture free from judgement

Data Sharing and Management Snafu in 3 Short Acts:



See: NYU Health Science Library, <https://www.youtube.com/watch?v=N2zK3sAtr-4>