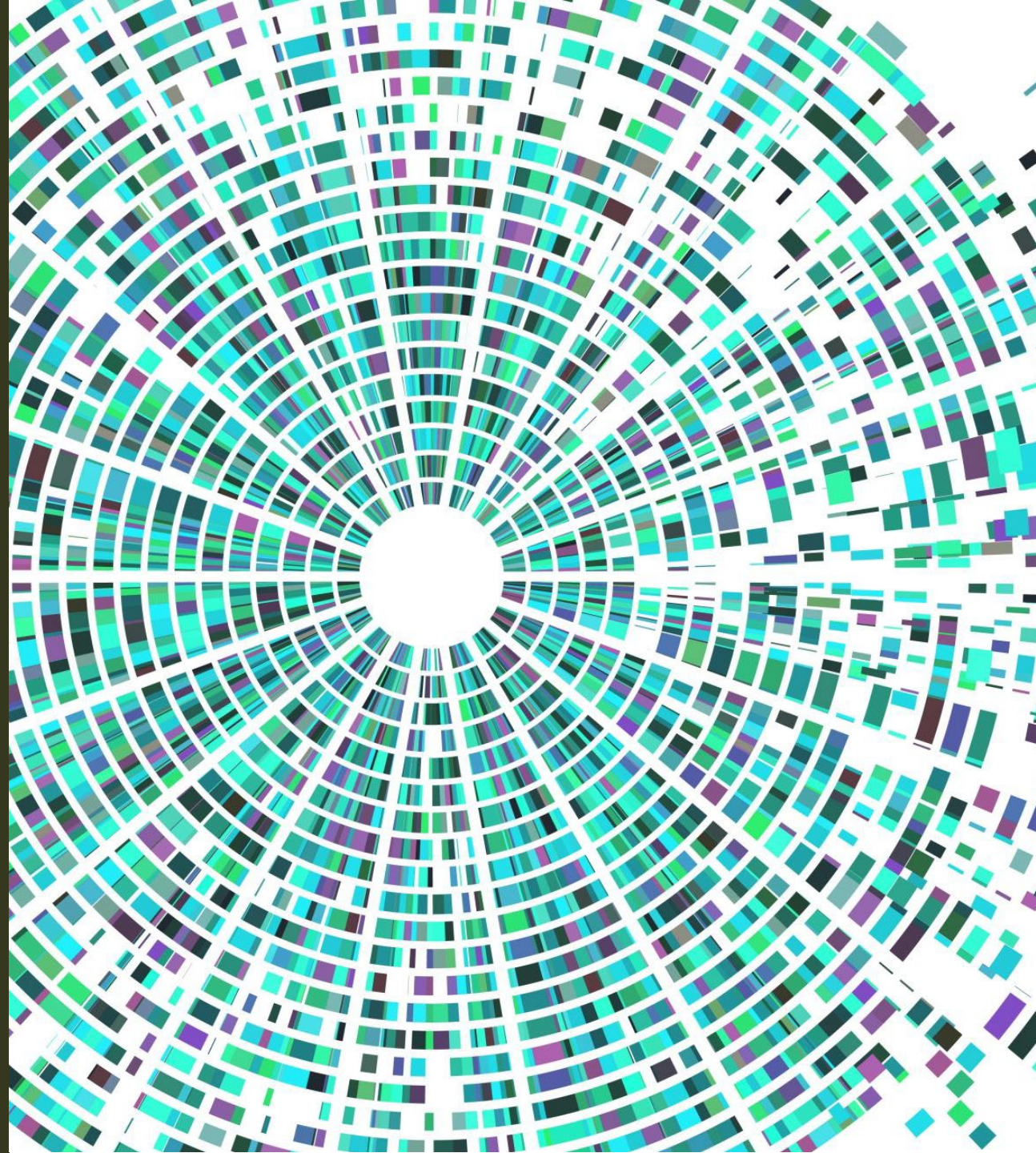# Data Governance & Stewardship 101

Lori Ann M. Schultz

Assistant Vice President, Research Intelligence

University of Arizona

What is Data Governance

A means to ensure:

- Data quality

- Data integrity

- Data security

- Data usability

For **all** of the data collected by an institution/organization.

What is included in Governance?

- People!

- Guidance on the accuracy, availability, use of data

- Specific information on access or restrictions on access to data

- Guidance on security to protect data and privacy

- Roles of those who interact with data

- Documentation, definitions

- Understanding business process

Why should we be concerned about it?

- We've moved beyond paper, which makes things more complicated

- Everyone "loves" data... ☺

- Trust is key

- Institutional policies

- Federal regulations

- Business sensitive information

- Privacy laws, regulations, expectations

WAYS TO SET UP QUALITY-FOCUSED DATA GOVERNANCE

Understand Your Data's Current State: Use data profiling to spot issues and set a benchmark for improvements.

Initiate Metrics and Monitoring: Develop KPIs and automated processes for tracking progress.

Standardize Data Quality: Create and communicate criteria for accuracy and completeness.

Foster Awareness through Training: Enhance awareness through educational initiatives.

Implement Quality Tools and Processes: Select appropriate tools and define workflows for regular data upkeep.

Encourage a Cycle of Continuous Improvement: Regularly review and adapt processes and tools.

Define Roles of Stewards and Owners: Assign roles to oversee data quality management.

Build a Collaborative Culture: Promote open dialogue and collaboration between teams.

Screenshot

From:
https://www.analytics8.com/blog/how-to-improve-data-quality-with-effective-data-governance/

What is Data Stewardship?

- Day-to-day oversight & management of data
  - Quality
  - Consistency
  - Accuracy

- Part of Data Governance and involves:
  - Business Processes & Systems
  - Policies

What are the rules/policies common to governance & stewardship?

- What people can access based on business need (roles/rights/permissions)

- What data is public on your organization's website

- Processes around data entry for any of your institutional systems

- Policies for data retention & destruction

- Training requirements for regulated data, appropriate use

Context matters…

- Governance process will vary!
  - Organization type: public, private, tribal, business, etc
  - Type of data handled
  - Risk tolerance of institution
  - Variant risk of data types
    - HR data
    - Financial data
    - Student data
    - Research data

Who is on your governance team?

Common roles:

- Data owners of your systems: research, financial, student, clinical, etc
- Data users: those who enter and use data
- Data Stewards: those responsible for day-to-day operation (like an eRA system)
- Data Custodians: Technical implementation, data storage, security
- Legal counsel
- IT / InfoSec
- Others?

Establishing (or learning about) data governance

- Clear roles & responsibilities: do you know who is responsible for the data in your eRA system? What kinds of people use the system and the resulting data?

- Policies and procedures

- Data quality expectations

- Accessibility & security – must be a balance

- Change management

- Scalability, monitoring, improvement

Why might governance fail?

- Most reasons are internal:

  - Lack of clear responsibilities
  - Lack of alignment with organizational goals
  - Institutional silos
  - Lack of sponsorship/accountability
  - Inadequate training/protection for regulated and sensitive data

Where does governance go wrong?

- It's not a "one & done" situation

- Framework for evolution is necessary

- Burden of constantly adding collection, but never reducing it

- Outdated rules for access/provisioning

Examples & Thoughts

- UArizona CRM: training and access considerations

- Proposal information: what is public?

- Employment application to first day of work

- Student admissions application to enrollment

- Considerations for generative AI tools

- Options to democratize data

Q&A