What do (some) data governors do all day?

Kathryn Flack Potts
Data Stewardship Coordinator
Stanford University

Alation, November 10, 2014

Data Governance

DG is the formalizing of behavior around the definition, production and usage of data to manage risk and improve the quality and usability of the selected data.*

Non-Invasive Data Governance, Robert Seiner, TDAN

* As the quality and the usability enables more efficient and effective operational processes and strategic decisions

Today's questions

Q1: What are we trying to do?

Q2: Who's involved?
What do they do (and why)?

Q3: How does technology fit in?

Premise: Everyone is different

Organizations: Needs, resources, culture

DG programs: Priorities, roles, processes

Data Governors, Steward(esse)s

$\neq$

Background: DG at Stanford

Primary driver:

Integrated BI reporting program

→ ~1 central FTE for DG

Tensions:

- Program vs. project orientation
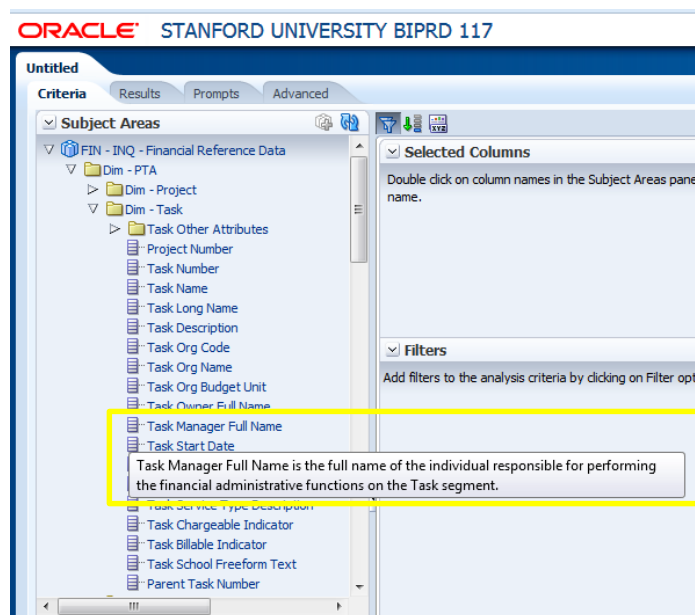- Central vs. local ownership
- Formal structure vs. DIY

Q1: What are we trying to do?

Stanford University Data Stewardship

SUDS-XXX mission and goals

Mission: In partnership with central units and the campus community, this committee is dedicated to providing the XXX metadata infrastructure to support improved decision-making, ensure information integrity, build data knowledge, and meet compliance requirements university-wide, enabling our partners to excel at teaching, learning and research.

Q2: Who's involved? What do they do?

Data Stewardship Coordinator (0.8 FTE)

Data Stewardship Analyst (0.5 FTE)

Chief Data Stewards: FIN, HR, STU, SPO
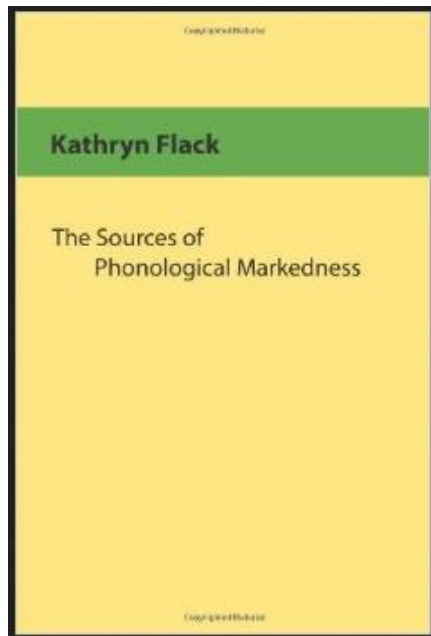
Chief Data Stewards:
Provide oversight, leadership, and signoff on all subject area stewardship activities.
Identify relevant stakeholders, on a topic-by-topic basis, within and across subject areas.
Participate in establishing stewardship standards and processes as part of SUDS-SC.
Serve as liaison with stakeholders and SMEs as needed.

…and whoever else we need to get things done

Interlude: KFP's backstory

STANFORD
UNIVERSITY

Institutional Research & Decision Support

Kathryn Flack

The Sources of
Phonological Markedness

Q2: Who's involved? What do they do?

Make SUDS happen

- Set project scope, goals, schedule
- Gather (the right) people
- Articulate immediate next steps
- Engage closely with content development
- Make sure things move forward
- **Make sure participants feel appreciated, and like things are moving forward**

Q2: Who's involved? What do they do?

Standards, best practices, documentation

Non-Standard Name	Standardized Name
Initial Submission →	Purchase Order Initial Submission Date

Required Term	Function	Example
Entity Name	What person, place, thing, or concept are we interested in?	Student, Employee, Award
Attribute Name	What property of this entity are we interested in?	Student Status, Employee Type, Award Purpose
Class Word	What type of data about this attribute is being presented?	Name, Indicator, Code

Q2: Who's involved? What do they do?

Side projects

(see http://irds.stanford.edu/maps)

Q2: Who <u>was</u> involved?
What did <u>they</u> do?

Data Governor #1

- Outreach, awareness
- "Here are things we could do. What would help?"

Q2: Who <u>was</u> involved?
What did <u>they</u> do?

DGer #1: "How about this stuff?"

- Data quality, profiling, root cause analysis, impact assessment

- Data lineage; source system overviews

- Policy: Ownership, access, usage

Q2: Who <u>was</u> involved?
What did <u>they</u> do?

DGer #1: <u>Data Governance Maturity Model</u>

Foundational	People	Policies	Capabilities
Awareness	② → ③ What awareness do people have about the their role within the data governance program?	② What awareness is there of data governance policies, standards and best practices?	① → ③ What awareness is there of data governance enabling capabilities that have been purchased or developed?
Formalization	② → ④ How developed is the data governance organization and which roles are filled to support data governance activities?	① → ② To what degree are data governance policies formally defined, implemented and enforced?	② → ③ How developed is the toolset that supports data governance activities and how consistently is that toolset utilized?
Metadata	① → ③ What level of cross functional participation is there in the development and maintenance of metadata?	② To what degree are metadata creation and maintenance policies formally defined, implemented and enforced?	② What capabilities are in place to actively manage metadata at various levels of maturity?

Project	People	Policies	Capabilities
Stewardship	② → ③ To what degree have stewardship roles been defined and filled?	① → ③ To what degree are stewardship policies defined, implemented and enforced?	② → ③ What capabilities are implemented to support the effective stewardship?
Data Quality	② → ③ To what degrees have data quality competencies developed?	① → ② To what degree are data quality policies defined, implemented and enforced?	① What capabilities are implemented to support the production and maintenance of high quality data?
Master Data	② To what degree has a formal master data management organization been developed and assigned consistent responsibilities across data domains?	② To what degree are master data policies defined, implemented and enforced?	② What capabilities are available and implemented to actively master and provision master data?

Q2: Who <u>was</u> involved?
What did <u>they</u> do?

Data Governor #2

- Top-down roles and structures

- **Metadata management tool selection**

We don't <u>need</u> specialized tools.

DG @ Stanford **Postdoctoral Scholar**

✏ Edit   ✉ Share   ➕ Add ▾   ⚙ Tools ▾

Added by Kathryn Flack Potts, last edited by Kathryn Flack Potts on Oct 30, 2013 (view change)

Definition: A Postdoctoral Scholar is a Non-Degree-Seeking Graduate Student pursuing advanced training beyond the doctoral level in preparation for an independent career.
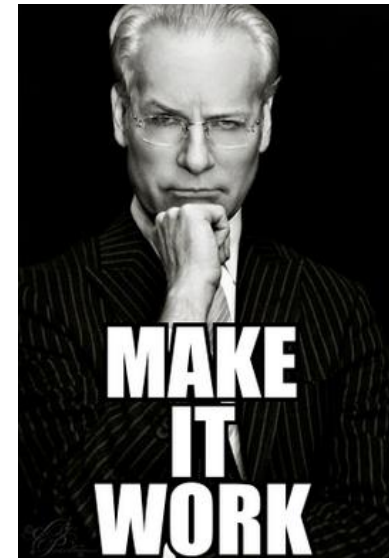
Aliases: Postdoc

Attributes:

Relationships: Postdoctoral Scholars are either Medical Fellows or Postdoctoral Research Fellows. An individual may only be enrolled as a Postdoctoral Scholar in one department at a time. An individual may be simultaneously enrolled as a Postdoctoral Scholar and in a Degree-granting program (typically a Master's or PhD program); in these cases, the individual is typically reported as a Degree-Seeking Student student rather than as a Postdoctoral Scholar.

Authoritative Source: PeopleSoft-Student is the authoritative source for active Postdoctoral Scholars on any given date.

Discrepancies: The alias "Postdoc" is commonly used both to refer to all Postdoctoral Scholars and to refer specifically to Postdoctoral Research Fellows. "Postdoc" is also used casually to refer to many different kinds of term-limited positions which are held by recent recipients of doctoral degrees, and which provide training or experience that make these individuals better prepared for careers, though these are not true Postdoctoral Scholar positions. IHUM instructor positions are an example of this kind of "quasi-postdoc".

While Postdoctoral Scholars are technically Non-Degree-Seeking Students, the Office of Postdoctoral Affairs tries not to refer to them as "students" or "graduate students" wherever possible, as they function very differently from typical Degree-Seeking Students.
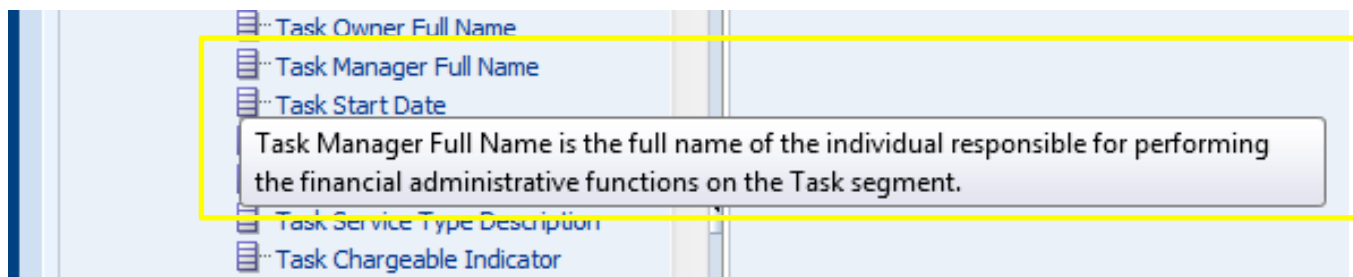
MAKE IT WORK

But they are really, really useful.

Q3: How does technology fit in?

Robust (yet flexible) metadata structures!

→ Decreased SME training

→ Increased consistency and quality

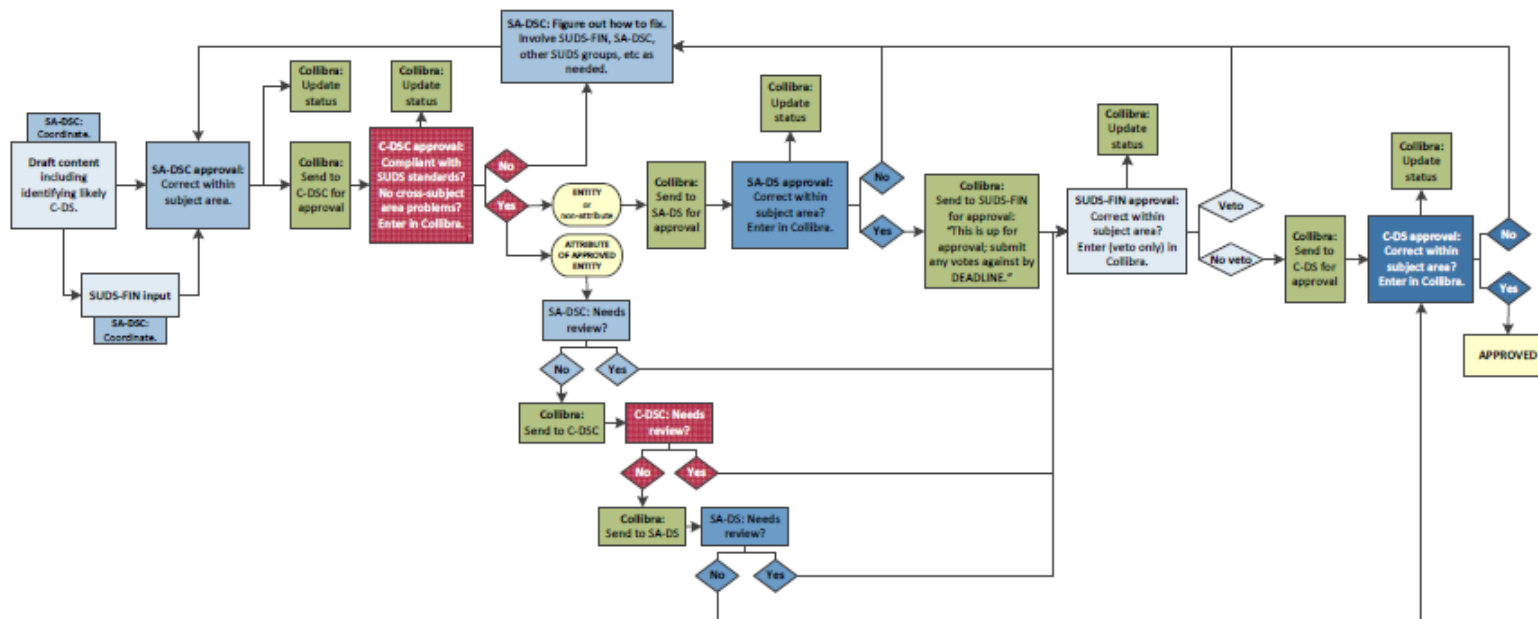**→ Massively easier to update, maintain, assess, and use**

- Task Owner Full Name
- Task Manager Full Name
- Task Start Date

Task Manager Full Name is the full name of the individual responsible for performing the financial administrative functions on the Task segment.

- Task Service Type Description
- Task Chargeable Indicator

→ Adaptable as needs evolve over time
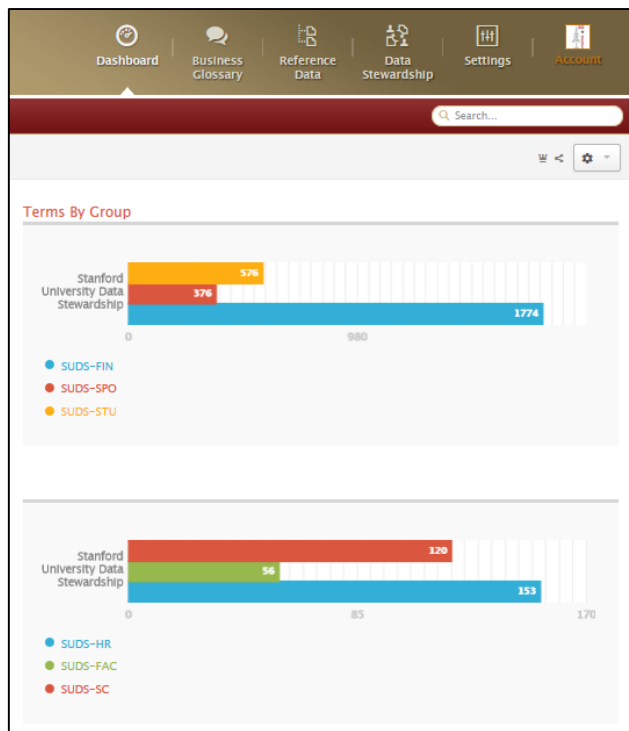
Q3: How does technology fit in?

Manual processes can be automated!

SUDS Metadata Approval Process (draft)

Q3: How does technology fit in?

Metadata is more visible and credible!

Broad engagement is less scary, more effective, and more rewarding.

Questions?