# Data Governance at Stanford:

## *The Stanford DG Maturity Model*

October 17, 2011

**Contents**


- Measuring Stanford's DG Maturity
- Qualitative and Quantitative Metrics
- Using a Maturity Model
- Short-term DG Activities


DG @ Stanford

## Measuring Stanford's DG Maturity

A maturity model is one of the most valuable tools available for planning and sustaining a new strategic program. Like the data governance (DG) program itself, the Stanford DG maturity model was designed around the unique goals, priorities and competencies of the institution. Many aspects of this model were adapted from other DG maturity models and customized for the program being developed here.

This maturity model is based on the structure of Stanford's DG program, focusing on both foundational and project aspects of DG. The foundational components of the maturity model (Awareness, Formalization and Metadata) focus on measuring core DG competencies and development of critical program resources. The project components (Stewardship, Data Quality and Master Data) measure how effectively DG concepts are applied in the course of funded projects.

> ***If you cannot measure it, you cannot improve it.***
>
> ***-Lord Kelvin***

Three dimensions (People, Policies and Capabilities) further subdivide each of the six maturity components, focusing on specific aspects of component maturation.

- People: Roles and organization structures.
- Polices: Development, auditing and enforcement of data policies, standards and best practices.
- Capabilities: Enabling technologies and techniques.

The model and descriptions of each metric can be found on the data governance website at http://dg.stanford.edu/wp-content/uploads/2011/11/StanfordDataGovernance-MaturityModel.pdf. It is imperative that the maturity model is finalized and adopted early in the rollout of the DG program and remains stable throughout its life. Thoughtful input from across the organization will help assure the model's long-term fitness. Please contact Matt Hoying hoying@stanford.edu with any recommendations for refinement.


**Contact Data Governance**

**Matt Hoying**

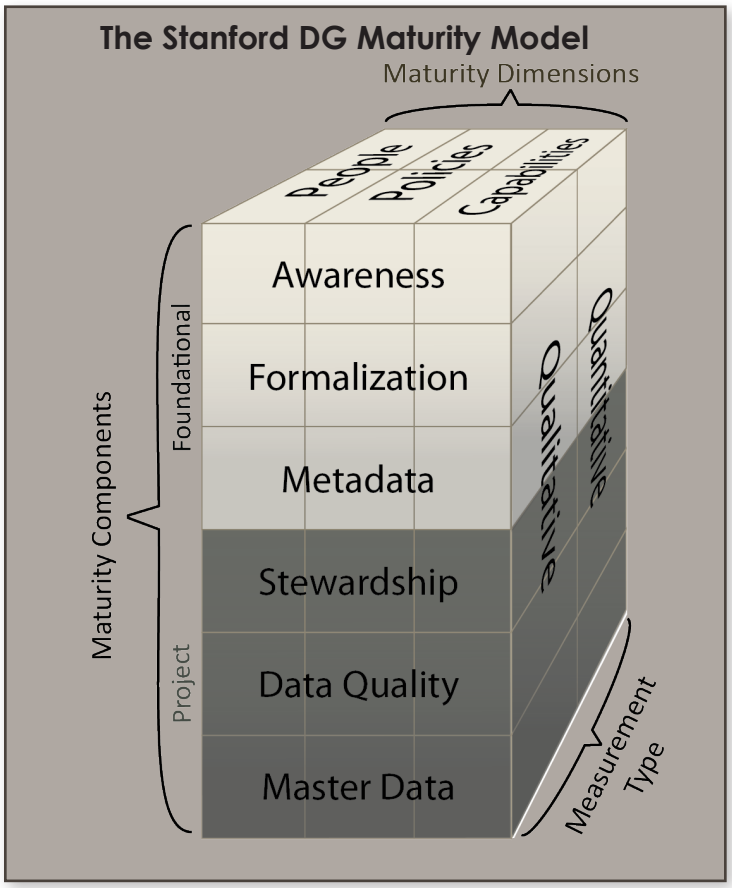
Data Governance Manager

hoying@stanford.edu

(650) 736-8192

http://dg.stanford.edu/

Institutional Research and Decision Support


## Qualitative and Quantitative Metrics

The Stanford data governance (DG) maturity model contains both qualitative and quantitative metrics to track the growth of the DG practice throughout the institution.

Qualitative aspects describe characteristics of the organization at various levels of maturity. Because these are inherently subjective, the model is enriched with quantitative metrics that count activities performed, program participants and artifacts developed.

Each component-dimension's qualitative scale ranges from level one, representing the initial state of a data governance program, to level five, representing the objective of DG in that area of focus. The quantitative metrics are numeric measures that become applicable at each level of maturity and may be used at all maturity levels moving forward.

Advancement through qualitative maturity levels can take place over a long time; quantitative metrics provide the abiltiy to monitor intra-stage growth through more granular measures.

# Using a Maturity Model

A maturity model is a tool that is used to develop, assess and refine an expansive program. Because measurement of performance simply through ROI or reduction of cost is inappropriate for data governance programs, another method must be constructed to assess effectiveness.

A significant benefit of utilizing a maturity model is that it can consistently measure the state of a program over time. A DG program crosses functional boundaries and has a lifespan measured in years rather than months. Stable metrics facilitate presentation of the DG program's accomplishments to the sponsors, ensuring the sustainability of the program, and demonstration to the participants that their efforts are driving organizational change.

The design of the maturity model also influences the strategic direction of the program. A maturity model is made up of levels describing possible states of the organization where the highest levels define a vision of the optimal future state.

Because the full implementation and maturation of a DG program is a multi-year effort, the intermediate maturity states can be used to construct a program roadmap. This outline describes a set of activities that will enable the institution to achieve their DG goals. The model not only faciliates assessment of the DG program, but also focuses attention on specific areas where actionable opportunities can be addressed rapidly.

## Short-term DG Activities

### In Progress

**Finalize DG maturity model and indicators**: Get feedback from BICC and SMEs on proposed maturity model's structure, content and metrics.

**Establish Stanford DG maturity baseline**: Use approved model to baseline current level of maturity and establish DG maturity goals.

**Develop metadata through HR Metrics project**: As a result of the HR Metrics Dashboard project, a team is working to construct definitions for terms that will apply across HR and provide the basis for the construction and auditing of institutional data definitions.

### Upcoming

**Develop institutional data subject areas**: A taxonomy of high level data subject areas (such as Research, Human Resources, Financial Resources, etc.) is key to assigning accountability and responsibility as well as designing an effective data stewardship organization.

**Define DG organizational structure**: Although there is overlap in responsibilities and content between a DG Executive Committee and the BICC, a dedicated Data Governance organization is necessary to keep focus on the continual development of data governance specific competencies.

**Define stewardship responsibilities**: The responsibilities assigned to data stewards must be carefully considered so that they are empowered to actively manage the definition and usage of their assigned data but not overwhelmed by the new responsibilities.

## Definitions

*Awareness*: The extent to which individuals within the organization have knowledge of the roles, rules and technologies associated with the data governance program.

*Data governance (DG)*: "The formalization of behavior around the definition, production and usage of data to manage risk and improve the quality and usability of the selected data." (Robert Seiner, TDAN)

*Data quality (DQ)*: "The continuous process for defining the parameters for specifying acceptable levels of data quality to meet business needs, and for ensuring that data quality meets these levels." (DMBOK, DAMA)

*Data entity*: "A concept that can, or does, take on one or more values" (ISO 21961:2003). A data entity is a collection of information that represents a class of thing (such as "student" or "course"). In a relational database, this usually maps to a table and each instance of that entity maps to rows within that table.

*Data profiling*: Gathering statistics and information about a selected set of data or data store. Profiling data is a key activity for data quality improvement. Profiling may be done at the element, entity or system level.

*Formalization*: "The extent to which work roles are structured in an organization and the activities of the employees are governed by rules and procedures." (http://www.businessdictionary.com/definition/formalization.html)

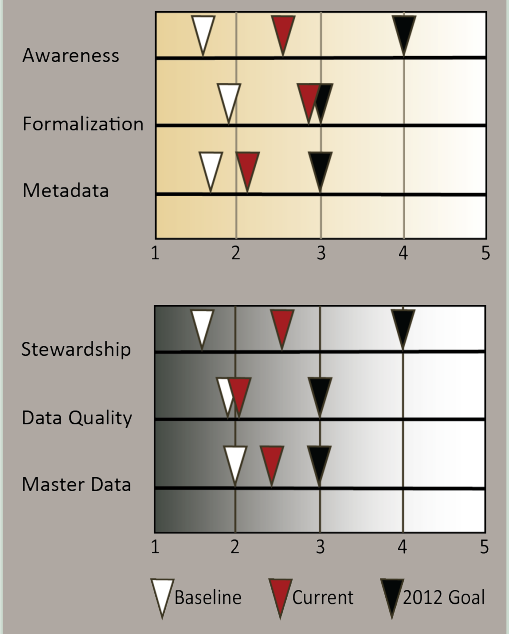
*Master data*: Business-critical data that is highly shared across the organization. Master data is often codified data, data describing the structure of the organization or key data entities (such as "employee" or "course").

*Metadata*: Data that 1) describes other data and IT assets (such as databases, tables and applications) by relating essential business and technical information and 2) facilitates the consistent understanding of the characteristics and usage of data. Technical metadata describes data elements and other IT assets as well as their use, representation, context and interrelations. Business metadata answers the who, what, where, when, why and how for users of the data and other IT assets.

*Stewardship (Data Stewardship)*: The formalization of accountability for the definition, usage and quality standards of specific data assets within a defined organizational scope.

*Subject area*: A grouping of information or concepts that represents a fundamental portion of the organization's data assets where all members are logically related.

### A Sample DG Maturity Scorecard




A well designed scorecard clearly shows the progress of an organization in terms of where it started and the short-term goals.

This document and the related presentation can be found online at: http://dg.stanford.edu/?tag=bicc