

Data Governance Maturity Model

Guiding Questions for each Component-Dimension

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

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

Data Governance Foundational Component Maturity

Maturity Level	Awareness			Formalization			Metadata		
	People	Policies	Capabilities	People	Policies	Capabilities	People	Policies	Capabilities
	1	Limited awareness of purpose or value of Data Governance program.	Most existing data policies are undocumented and there may be inconsistent understanding of data policies within a department.	There is little awareness of data governance capabilities and technologies.	There are no defined roles related to data governance.	No formal data governance policies.	Classes of data governance capabilities are not defined.	Limited understanding of types and value of metadata.	No metadata related policies
2	Executives are aware of existence of program. Little knowledge of program outside upper management.	Existing policies are documented but not consistently maintained, available or consistent between departments.	A small subset of the organization understands the general classes of data governance capabilities and technologies.	Data governance roles and responsibilities have been defined and vetted with program sponsors.	High-level data governance meta-policies are defined and distributed.	Classes of data governance capabilities are defined and home-grown technical solutions are used within some institutional functions.	AS roles responsible for production of technical metadata on structured data are defined during system design.	Metadata best practices are produced and made available. Most best practices are focused on the metadata associated with structured data.	Metadata templates are adopted to provide some consistency in content and format of captured metadata. Metadata is consolidated and available from a single portal. Capabilities focus on capture of metadata of structured content.
3	Executives understand how program benefits/impacts their portion of the organization, knowledge workers are aware of program. Executives actively promote program within their groups.	Common institutional data policies are documented and available through a common portal. Most stakeholders are aware of existence of data policies that may impact them.	A small subset of the organization is aware of the specific data governance capabilities that are available at Stanford.	Some roles are filled to support data governance needs and participants clearly understand responsibilities associated with their roles.	Data policies around the governance of specific data are defined and distributed as best practices.	Home-grown technical solutions are adopted as best practices for some classes of capabilities and made available throughout the institution.	The responsibility for developing institutional business definitions and storing them in a central repository is assigned to and continually performed by subject matter experts.	Policies requiring the development of new metadata as part of system development (usually focused on structured data) are adopted as official Stanford data policies.	The collection of metadata on structured content is automated and scheduled extracts are performed for selected systems.
4	Executives understand long-term program strategy and their part in it. Knowledge workers understand how the program impacts/benefits their portion of the organization. Executives actively promote program beyond the immediate group.	All data policies are available through a common portal and stakeholders are actively notified whenever policies are added, updated or modified.	A targeted audience has been identified and a significant portion of that audience is aware of the data governance capabilities that are available at Stanford.	Data governance roles are organized into reusable schemas which are designed to support specific data and functional characteristics. There is broad (but inconsistent) participation in Data Governance Organization.	Data policies become official Stanford data policies and compliance with approved data policies is audited.	All defined classes data governance capabilities have an available solution.	Metadata collection/validation responsibilities assigned to named individuals for all projects.	Policies requiring the regular auditing of metadata in specified systems are adopted as official Stanford data policies and metadata development as part of system development is enforced.	A centralized metadata store becomes the primary location for all institutional metadata. Metadata is automatically collected from most RDBMS and vendor packaged systems.
5	Both executives and knowledge workers understand their role in the long-term evolution of the program. Knowledge workers actively promote program.	A history of all data policies are maintained through a common portal and all stakeholders are made part of the policy development process through online collaborative tools.	A significant portion of the targeted audience understands how to utilize relevant data governance capabilities that are available at Stanford.	Data governance organizational schemas are filled as defined, meet regularly and document activities.	Compliance with official Stanford data policies is actively enforced by a governing body.	All defined classes of data governance capabilities are mandatory for assigned systems or critical data.	A dedicated metadata management group is created to strategically advance metadata capabilities and more effectively leverage existing metadata.	Metadata policy covers both structured and unstructured (non-tabular) data and is enforced.	A metadata solution provides a single point of access to federated metadata resources including both structured and unstructured data.

Data Governance Project Component Maturity

Maturity Level

	Stewardship		
	People	Policies	Capabilities
1	Almost no well defined data governance or stewardship roles or responsibilities. Data requirements are driven by the application development team.	Limited stewardship policies documented.	Limited stewardship capabilities are available.
2	Business analysts drive data requirements during design process. Definition of stewardship roles and responsibilities is limited.	Policies around stewardship defined within a functional area or subject area.	A centralized location exists for consolidation of and/or access to stewardship related documentation.
3	All stewardship roles and structures are defined and filled but are still functionally siloed.	Stewardship policies are consistent between functions and subject areas.	Workflow capabilities are implemented for the vetting and approval of institutional definition, business metadata and approval of other stewardship related documentation.
4	The stewardship structures include representatives from multiple business functions.	Stewardship teams self-audit compliance with policies.	Stewardship dashboards report data quality levels and data exceptions to support the auditing of stewardship effectiveness.
5	The stewardship board includes representatives from all relevant institutional functions including AS.	Compliance with stewardship policies are enforced for key institutional data.	A common stewardship dashboard enables managed issue remediation as part of data quality reporting and data exception reporting.

	Data Quality		
	People	Policies	Capabilities
1	Individuals perform ad hoc data quality efforts as needed and manually fix data identified data issues. Identification of data issues is based off of its usability for a specific business task.	Data quality efforts are infrequent and driven by specific business needs. These efforts are usually large one-time data cleansing efforts.	Data quality is done on ad hoc basis usually using SQL and Excel.
2	A small group of individuals are trained in and perform profiling to assess data quality of existing systems to establish a baseline or justify a data quality project. Down stream usage of the data is considered in issue identification process.	Best practices have been defined for some data quality related activities and followed inconsistently.	Basic data profiling tools are adopted and available for use anywhere in the system development lifecycle.
3	People are assigned to assess and ensure data quality within the scope of each project.	Profiling and development of data quality standards are adopted as part of the standard application development lifecycle and become scheduled activities on project plans.	Data quality reporting capabilities are implemented and available to any Stanford system.
4	Data quality experts are identified throughout the institution and are engaged in all data quality improvement projects.	Data quality best practices are adopted as official Stanford data policies.	Data quality issue remediation is integrated into quality reporting platform.
5	A data quality competency center is funded and charged with continually assessing and improving data quality outside of the system development lifecycle.	Compliance with official Stanford data quality policies is tracked and reported on centrally.	Data quality remediation is implemented on both data at rest (in databases) and data in flight (in ETL and as messages between systems).

	Master data		
	People	Policies	Capabilities
1	Inconsistent understanding of concepts and benefits of Master Data Management.	No formal policies defining what data is considered institutional master data.	There is limited management of master data.
2	Stakeholders for specific master data domains are identified and consulted to develop basic definition and model of master data.	Institutional master data domains are defined and the systems storing master data and is documented. Usage of master data in these systems is actively being documented.	Master data are identified and manually managed and provisioned via extracts, file transfers or manual uploads.
3	Owners of institutional master data are identified and drive resolution of various perspectives of master data. Owners establish and run master data boards to support maintenance and data issue mediation.	Institutional master data perspectives resolved and documented.	Master data are provisioned through services but management capabilities are still largely manual.
4	Master Data Management boards take responsibility for reviewing the use of their master data in the application development process.	Compliance with master data usage policies and standards is enforced. Synchronization frequency with master data hub at system owner's discretion.	Multiple single domain master data hubs handle provisioning and management of master data.
5	Master Data Management boards take responsibility for enforcing master data policies around their own master data across the institution.	Compliance with master data synchronization policy is enforced.	Multidomain master data hub handles all provisioning and management of master data.

Foundational Components

	People		Policies		Capabilities		
	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	
	Awareness	1	Limited awareness of purpose or value of Data Governance program.		Most existing data policies are undocumented and there may be inconsistent understanding of data policies within a department.		There is little awareness of data governance capabilities and technologies.
2		Executives are aware of existence of program. Little knowledge of program outside upper management.	Training Sessions*attendees	Existing policies are documented but not consistently maintained, available or consistent between departments.	Policies documented by functional area, business subject area.	A small subset of the organization understands the general classes of data governance capabilities and technologies.	Training Sessions on DG capabilities and technologies
3		Executives understand how program benefits/impacts their portion of the organization, knowledge workers are aware of program. Executives actively promote program within their groups.	Newsletters*recipients	Common institutional data policies are documented and available through a common portal. Most stakeholders are aware of existence of data policies that may impact them.	Hits on Policy Management Content, Unique visitors	A small subset of the organization is aware of the specific data governance capabilities that are available at Stanford.	
4		Executives understand long-term program strategy and their part in it. Knowledge workers understand how the program impacts/benefits their portion of the organization. Executives actively promote program beyond the immediate group.	Hits on DG website, Unique visitors on DG website	All data policies are available though a common portal and stakeholders are actively notified whenever policies are added, updated or modified.	Number of stakeholders in RACI matrices by subject area, functional area	A targeted audience has been identified and a significant portion of that audience is aware of the data governance capabilities that are available at Stanford.	
5		Both executives and knowledge workers understand their role in the long-term evolution of the program. Knowledge workers actively promote program.		A history of all data policies are maintained through a common portal and all stakeholders are made part of the policy development process through online collaborative tools.	Non-board participants in policy development	A significant portion of the targeted audience understands how to utilize relevant data governance capabilities that are available at Stanford.	Training Sessions on usage of DG technologies and capabilities (person*tech trained)

	People		Policies		Capabilities		
	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	
	Formalization	1	There are no defined roles related to data governance.		No formal data governance policies.		Classes of data governance capabilities are not defined.
2		Data governance roles and responsibilities have been defined and vetted with program sponsors.		High-level data governance meta-policies are defined and distributed.	Meta-policies defined, documented and approved	Classes of data governance capabilities are defined and home-grown technical solutions are used within some institutional functions.	Data governance capabilities with solutions by functional org. Reuse of technical solutions by functional org.
3		Some roles are filled to support data governance needs and participants clearly understand responsibilities associated with their roles.	Participants in approved roles.	Data policies around the governance of specific data are defined and distributed as best practices.	Best Practices/Standards/Policies identified, documented and approved	Home-grown technical solutions are adopted as best practices for some classes of capabilities and made available throughout the institution.	Capabilities approved as Stanford recommended solutions.
4		Data governance roles are organized into reusable schemas which are designed to support specific data and functional characteristics. There is broad (but inconsistent) participation in Data Governance Organization.	Boards in compliance with defined schemas, % roles filled.	Data policies become official Stanford data policies and compliance with approved data policies is audited.	Official Data policies approved, data policies with audit	All defined classes data governance capabilities have an available solution.	Usage of standard solutions by project. Uses of non-standard solutions by project
5		Data governance organizational schemas are filled as defined, meet regularly and document activities.	Board meetings to plan, Minutes produced.	Compliance with official Stanford data policies is actively enforced by a governing body.	(Lower is better) Exceptions to official data policies.	All defined classes of data governance capabilities are mandatory for assigned systems or critical data.	(Lower is better): Uses of non-standard solutions by project, (Lower is better) no use of solution by project.

	People		Policies		Capabilities		
	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	
	Metadata	1	Limited understanding of types and value of metadata.		No metadata related policies		Metadata is inconsistently collected and rarely consolidated outside of project artifacts.
2		AS roles responsible for production of technical metadata on structured data are defined during system design.	(Lower is better) Projects without documented technical metadata.	Metadata best practices are produced and made available. Most best practices are focused on the metadata associated with structured data.	Best Practices identified, in progress, approved.	Metadata templates are adopted to provide some consistency in content and format of captured metadata. Metadata is consolidated and available from a single portal. Capabilities focus on capture of metadata of structured content.	Metadata entities in portal. Edits by users to metadata
3		The responsibility for developing institutional business definitions and storing them in a central repository is assigned to and continually performed by subject matter experts.	Unique individuals creating/updating metadata. Qualitative rating of metadata.	Policies requiring the development of new metadata as part of system development (usually focused on structured data) are adopted as official Stanford data policies.	New Metadata entities/elements by project, metadata reuse. (Lower is better) projects without metadata policy.	The collection of metadata on structured content is automated and scheduled extracts are performed for selected systems.	Systems with automatic collection of metadata.
4		Metadata collection/validation responsibilities assigned to named individuals for all projects.	Projects with metadata responsibility assignment	Policies requiring the regular auditing of metadata in specified systems are adopted as official Stanford data policies and metadata development as part of system development is enforced.	Systems with audits in place. Compliance with policy.	A centralized metadata store becomes the primary location for all institutional metadata. Metadata is automatically collected from most RDBMS and vendor packaged systems.	(Lower is better) Systems not loading to metadata repository.
5		A dedicated metadata management group is created to strategically advance metadata capabilities and more effectively leverage existing metadata.	ROI of Metadata Competency Center	Metadata policy covers both structured and unstructured (non-tabular) data and is enforced.	Structured Policies/Systems in compliance, Unstructured Policies/Objects in compliance.	A metadata solution provides a single point of access to federated metadata resources including both structured and unstructured data.	Unstructured objects linked to metadata repository. (Lower is better) systems out of compliance with load SLA.

Project Components

	People		Policies		Capabilities		
	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	
Stewardship	1	Almost no well defined data governance or stewardship roles or responsibilities. Data requirements are driven by the application development team.		Limited stewardship policies documented.		Limited stewardship capabilities are available.	
	2	Business analysts drive data requirements during design process. Definition of stewardship roles and responsibilities is limited.	Projects with explicit Data Design.	Policies around stewardship defined within a functional area or subject area.	Functional areas with policy, Functional Data Entities with policy	A centralized location exists for consolidation of and/or access to stewardship related documentation.	Count of policies (by status) in registry
	3	All stewardship roles and structures are defined and filled but are still functionally siloed.	Stewards, Participants in Stewardship boards, Stewardship board meetings.	Stewardship policies are consistent between functions and subject areas.	Institutional Data Entities with policy	Workflow capabilities are implemented for the vetting and approval of institutional definition, business metadata and approval of other stewardship related documentation.	Institutional Definitions through process (completed, in progress)
	4	The stewardship structures include representatives from multiple business functions.	Functional Areas Represented on Stewardship Boards	Stewardship teams self-audit compliance with policies.	Audits, Audit Compliance	Stewardship dashboards report data quality levels and data exceptions to support the auditing of stewardship effectiveness.	Dashboards by function, subject area. Qualitative score on dashboard.
	5	The stewardship board includes representatives from all relevant institutional functions including AS.	Boards with AS and Business Representation	Compliance with stewardship policies are enforced for key institutional data.	(Lower is better) Key Institutional Data without stewardship policies	A common stewardship dashboard enables managed issue remediation as part of data quality reporting and data exception reporting.	Issues reported, Issues resolved, Time to resolution.

	People		Policies		Capabilities		
	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	
Data Quality	1	Individuals perform ad hoc data quality efforts as needed and manually fix data identified data issues. Identification of data issues is based off of its usability for a specific business task.	*data quality implies quality in terms of formally defined definition of fit-for-use data.	Data quality efforts are infrequent and driven by specific business needs. These efforts are usually large one-time data cleansing efforts.	Data Cleansing Efforts identified, in progress, completed.	Data quality is done on ad hoc basis usually using SQL and Excel.	
	2	A small group of individuals are trained in and perform profiling to assess data quality of existing systems to establish a baseline or justify a data quality project. Down stream usage of the data is considered in issue identification process.	Individuals trained in profiling, systems profiled, tables profiled, elements profiled. Profiles resulting in recommendations, recommendations spawning projects.	Best practices have been defined for some data quality related activities and followed inconsistently.	Data Quality Best Practices defined.	Basic data profiling tools are adopted and available for use anywhere in the system development lifecycle.	Data Profiles by system, functional area. Rows profiled.
	3	People are assigned to assess and ensure data quality within the scope of each project.	Projects with DQ roles assigned. "DQ fixes" at project level. Issues documented and approved.	Profiling and development of data quality standards are adopted as part of the standard application development lifecycle and become scheduled activities on project plans.	(Lower is better) Application development projects without profiling effort.	Data quality reporting capabilities are implemented and available to any Stanford system.	Systems with data quality reporting, approved elements reported on. Raw Quality Metrics.
	4	Data quality experts are identified throughout the institution and are engaged in all data quality improvement projects.	Systems analyzed, tables analyzed, elements analyzed. Recommendations proposed, Recommendations spawning DQ remediation.	Data quality best practices are adopted as official Stanford data policies.	Approved Stanford DQ Policies, data quality policies with audit	Data quality issue remediation is integrated into quality reporting platform.	Systems with data quality remediation functionality. Issues resolved.
	5	A data quality competency center is funded and charged with continually assessing and improving data quality outside of the system development lifecycle.	ROI of DQCC. System Team endorsements.	Compliance with official Stanford data quality policies is tracked and reported on centrally.	(Lower is better) Exceptions to official data quality policies.	Data quality remediation is implemented on both data at rest (in databases) and data in flight (in ETL and as messages between systems).	(Lower is better) Systems without DQ reporting, remediation. Interfaces without reporting, remediation.

	People		Policies		Capabilities		
	Qualitative	Quantitative	Qualitative	Quantitative	Qualitative	Quantitative	
Master Data	1	Inconsistent understanding of concepts and benefits of Master Data Management.		No formal policies defining what data is considered institutional master data.		There is limited management of master data.	
	2	Stakeholders for specific master data domains are identified and consulted to develop basic definition and model of master data.	Stakeholders identified, stakeholders' agreement.	Institutional master data domains are defined and the systems storing master data and is documented. Usage of master data in these systems is actively being documented.	Master Data Entities Identified. Functions consulted. Perspectives Identified.	Master data are identified and manually managed and provisioned via extracts, file transfers or manual uploads.	Systems using master data by transport method
	3	Owners of institutional master data are identified and drive resolution of various perspectives of master data. Owners establish and run master data boards to support maintenance and data issue mediation.	Approved owners, stakeholders with input.	Institutional master data perspectives resolved and documented.	Master Data Models approved. (Lower is better) distinct perspectives of master data entities.	Master data are provisioned through services but management capabilities are still largely manual.	Systems using master data via services.
	4	Master Data Management boards take responsibility for reviewing the use of their master data in the application development process.	Boards taking review responsibility.	Compliance with master data usage policies and standards is enforced. Synchronization frequency with master data hub at system owner's discretion.	Results of audit.	Multiple single domain master data hubs handle provisioning and management of master data.	Master Data Hubs. Master data hub capability score.
	5	Master Data Management boards take responsibility for enforcing master data policies around their own master data across the institution.	Boards taking enforcement responsibility.	Compliance with master data synchronization policy is enforced.	Results of audit.	Multidomain master data hub handles all provisioning and management of master data.	(Lower is better) Master data hubs, master data hub score.