Data Stewards Quarterly Meeting

Office of Data Governance and Data Governance Support Services & MDM

April 23, 2020
• Introduction  Adrian
• Office of Data Governance update  Eduardo
• DG Roles & Responsibilities  Oksana
• AIDA Updates  Adrian
• ODG Project Updates  Sara, Tammana
• Collibra DGC Updates  Chrissy
• Announcements & Events
• Q&A
Office of Data Governance updates
Data Governance Roles & Responsibilities
Data Stewards

- Typically operational managers who are subject matter experts of data in a functional/business area (student affairs, HR, etc.)
- Appointed by their respective Data Trustees
- Manage business processes, define domain specific business data definition, establish the business rules for the transaction systems
- First line of contact to resolve DQ issues, data process issues, and symptoms of DQ issues
- Identify and document data quality issues and business solutions
- Ensure adherence and alignment to NYU data standards and policies
- Provide business knowledge and guidance (from business process perspective) to Data Custodians
- Collaborate with other Central and School Stewards
- Responsible for business data assets
Data Custodians

- Responsible for technical data assets, define and manage technical metadata, and technical definition of the critical data element and link to the business metadata.
- Oversee the safe transport and storage of data, establish and maintain the underlying infrastructure, and perform activities required to keep the data intact and available to users.
- Information technology experts assigned to each transactional and reporting system maintaining administrative data.
- Enter and maintain correct data values in systems to support University operations, ensure implementation of business rules to production data.
- Monitor and assure data quality in support of / under the guidance of the Data Steward.
- Identify need for new data asset definitions and DQ issues; escalate to Data Stewards and/or Office of Data Governance / Data Governance Support Services.
- Uses Security to reduce risk through access management and technical controls.
- Develop automated processes which identify erroneous, inconsistent, or missing data.
- Resolve technical data issues.
Data Trustees

- Senior leaders in operational areas accountable for maintaining the content of transactional systems
- Serve as the first escalation for policy / problem resolutions from Data Stewards
- Support continuous improvement of University data governance processes and systems
- Review, approve, and implement policy
- Identify and assign Data Stewards
Data Governance

• Data Governance includes two teams, the Office of Data Governance (ODG) and Data Governance Support Services (DGSS).

  • **ODG sits within the Office of Institutional Research and Data Integrity (OIRDI) and focuses on:**
    • Developing and facilitating Data Governance, Data Quality, and Data Risk Control processes and areas
    • Developing and implementing data governance standards across the organization
    • Developing data governance policies in partnership with Data Trustees, Data Stewards, Data Governance Support Services, the Office of Compliance, and the Office of the General Council
    • Defining data governance roles and responsibilities, and managing communications across the organization

  • **DGSS sits within NYUIT and focuses on the technical side of data governance issues and focuses on:**
    • Defining and implementing data governance and data quality processes and improvements
    • Defining, developing and facilitating metadata management and data management processes
    • Ensuring NYU's technology investments support data strategy
    • Integrating data governance across initiatives and governance bodies across NYU

• Partners to execute the strategic direction and facilitate processes to achieve NYU’s goals
• Responsible for the data management framework
• Takes joint actions and makes recommendations on wide-ranging Data Governance issues
• Proactively and strategically coordinates with ongoing NYU data and analytics initiatives via direct partnership between IR and IT, working closely with Data Trustees, Stewards, Custodians, and Decision Support Group
Data Quality Issues
• Data is not fit for its intended use
• Missing, incorrect, inconsistent, conflicting, redundant, non-compliant

Operational Issue
• Issue in the quality of data resulting from operational inefficiencies or limitations

Process Issue
• Issue in the quality of data resulting from a lack of controls and checkpoints in data processes

Entry Issue
• Issues in the quality of data resulting from incorrectly entered information at the source

Symptom of DQ Issue
• We need to be able to distinguish what is a symptom versus a cause of data quality issues. Symptoms may result from a lack of controls and checkpoints at the point of entry, within processes, or during ETL or analysis.
Feedback and Discussion

- “Data Domain Trustee” vs “Data Trustee” – subject to formal approval
- Accessibility of Data Trustees – may be dependent on team
  - *Data Trustees serve as the first escalation for policy / problem resolutions from Data Stewards*
- Central versus School Data Stewards – focused on Central right now
- DG Teams consulting Custodians – Custodians should be included
  - *Data Governance proactively and strategically coordinates with ongoing NYU data and analytics initiatives via direct partnership between IR and IT, working closely with Data Trustees, Stewards, Custodians, and Decision Support Group*

Example

<table>
<thead>
<tr>
<th>Business Area</th>
<th>Data Trustee</th>
<th>Data Steward</th>
<th>Data Custodian</th>
</tr>
</thead>
<tbody>
<tr>
<td>University Development and Alumni Relations</td>
<td>Development and Alumni Relations (Advance)</td>
<td>Scott Jenks</td>
<td>Lisa Murphy</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Ping Gallivan</td>
</tr>
</tbody>
</table>
Please respond to survey if you have not already!

Link: [DG Roles & Responsibilities Questionnaire](#)
AIDA Updates
AIDA (Data Warehouse)

- **Recently Done**
  - BI Developers, Data Consumers can answer questions based on newly-available data:
    - Events and Activities (Center for Student Life: OrgSync, Event CheckIn)
    - Housing & Residential Life
    - Athletics (Varsity)
  - Reliance on Registry is being reduced as critical data is made available through AIDA:
    - Emma (Phase I)
  - Utilizing *Data Vault 2.0* methodology

- **Doing**
  - BI Developers, Data Consumers can answer questions around building utilization
    - Master building codes provide consistency in reporting across systems and cases
    - Transactions in operational systems will benefit from mastered data (e.g., SMS)
  - Reliance on Registry is being reduced as more data is available through AIDA:
    - Emma (Phase II)
  - Maturing the use of *Data Vault 2.0* with training, automation, and updated tools
Student Data Functional Architecture

Space Data Functional Architecture

Note: Topics in blue indicate current work in progress.
AIDA (Collibra)

● Recently Done
  • Developers, Data Consumers can find business terms & definitions in Business Glossary;
  • Data Stewards can direct inquiries on definitions to this standardized, enterprise resource available for an initial (and growing) set of business areas

● Doing
  • Legacy NYU Business Glossary sunset; terms & definitions in Collibra Business Glossary
  • Data Stewards can develop, govern terms & definitions using Collibra (i.e., Workflow)
  • Data Engineers, BI Developers can find, understand technical metadata
  • Reduce reliance on legacy software (i.e., streamline, savings)

● Highlights
  • Online training, certification available via Collibra University

● Questions?
  • Contact data.governance.support@nyu.edu
ODG Project Updates
**Objective:** Strengthen the quality and reliability of NYU’s census data by streamlining and making more efficient the process by which we validate this data.

### Current Issues:
- Short timeframe for error resolution
- Many manual steps, leaving potential for human error
- More we want to be validating
- Process involves “middle-man”

### Resources:
- Data Governance Support
- Office of Institutional Research
- Office of the University Registrar
- NYU IT (IT Ops, EDM, UDW+ ETL, SIS ETL)

### Risk:
- Could lead to incorrect internal and external reporting
- Could lead to incorrect billing in the Bursar Office

### Stakeholders:
- Registrar
- OIR
- Bursar Office
- Budget Office
- UDW+
- Reporting
- API
Objective: Strengthen the quality and reliability of NYU’s census data by streamlining and making more efficient the process by which we validate this data.

<table>
<thead>
<tr>
<th>Phases</th>
<th>Key Activities</th>
<th>Status</th>
<th>Progress Updates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery Phase I: Map Current State</td>
<td>• Map current process&lt;br&gt;• Document current DQ rules&lt;br&gt;• Group DQ rules by DQ dimensions</td>
<td>In Progress; On Schedule</td>
<td>• Project Team Kick Off 3/24&lt;br&gt;• High level process mapped&lt;br&gt;• Working with Registrar to finish documenting DQ rule logic</td>
</tr>
<tr>
<td>Discovery Phase II: Determine Future State</td>
<td>• Determine technology for future state solution</td>
<td>In Progress; On Schedule</td>
<td>• Met with various IT resources&lt;br&gt;• Reviewed possible prototype process</td>
</tr>
<tr>
<td>Implementation: Build Future State</td>
<td>• Build future state&lt;br&gt;• Reliability testing</td>
<td>Not Started; On Schedule</td>
<td>Not started</td>
</tr>
<tr>
<td>Post-Implementation: Pilot &amp; Train</td>
<td>• Pilot future state with Summer Census&lt;br&gt;• Change Mgmt &amp; Education</td>
<td>Not Started; On Schedule</td>
<td>Not started</td>
</tr>
</tbody>
</table>
Census Data Validation Process Improvement initiative: Project Team

Office of Data Governance
- Sara Dicker, Data Governance Analyst
- Oksana Kinker, Director of Data Governance and External Reporting

Data Governance Support
- Adrian Kulpa, Senior Director, Data Governance Support Services and Master Data Management
- Huey-Chih Lee, Project Manager

Office of Institutional Research and Data Integrity
- David Vintinner, Executive Director of Institutional Research

Office of the University Registrar
- Al Gentile, Associate Registrar
- Laurie Giannisis, Associate Registrar
- Tyrell Davis, Associate Registrar

NYU IT Operational Support
- Felisa Guzman, Manager of Data Analysis
- Jenny Alulema, Senior Systems Analyst

API
- Satya Kunta, Program Director, Technology Architecture, Data Management & Reporting

ETL Teams
- Courtney Marcus, Data Services Project Director
- Ramachandran Boopalam, Data Services Project Director

UDW+
- Andy Maliszewski, EDM Program Director

Reporting
- Addis Crooks-Jones, Director, Information and Analytics Services

Budget Office
- Cesar Villatoro, Senior Financial Analyst

Bursar Office
- TBD
Objective: To standardize academic and administrative names & codes across NYU enterprises effecting reports that determine data access and engage with senior leadership.

<table>
<thead>
<tr>
<th>Current Issues:</th>
<th>Resources (Technical &amp; Business):</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Historical logic is inconsistent across systems / schools / admins</td>
<td>• Data Governance Support Services</td>
</tr>
<tr>
<td>• No documentation for historical process.</td>
<td>• Work Day (NYUIT)</td>
</tr>
<tr>
<td>• Data Governance Support Services</td>
<td>• Application Programming Interface</td>
</tr>
<tr>
<td>Risk:</td>
<td>• Registrar</td>
</tr>
<tr>
<td>• Reporting errors due to incorrect manual data entry</td>
<td>• Budget</td>
</tr>
<tr>
<td>• Limited risk control – no standards for tracking changes</td>
<td></td>
</tr>
</tbody>
</table>

Stakeholders (Report Owners):

- Enterprise Data Management
- Decision Support Group

<table>
<thead>
<tr>
<th>Short Term</th>
<th>Long Term</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Analyze and improve logic for the RDC process if possible</td>
<td>• Identify scope for standardizing academic and administrative names &amp; codes across university reports</td>
</tr>
<tr>
<td>• Address escalations from end users of the RDC (Teams: API, EDM, and DGSS)</td>
<td>• Partner with DGSS to automate the RDC process.</td>
</tr>
<tr>
<td>• Implement change management (roles &amp; responsibilities).</td>
<td>• Finalize the reference list of systems, reports, and roles.</td>
</tr>
<tr>
<td>• Create a reference list of systems, reports, roles.</td>
<td>• Create standards, procedures, workflows, and etc.</td>
</tr>
</tbody>
</table>
Objective: To standardize academic and administrative names & codes across NYU enterprises effecting reports that determine data access and engage with senior leadership.

<table>
<thead>
<tr>
<th>Phases</th>
<th>Key Activities</th>
<th>Status</th>
<th>Progress Updates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Discovery Phase I:</td>
<td>[Map Current State](Est: 3wks, 03/16 – 04/03)</td>
<td>Complete; On Schedule</td>
<td>• Project Team Kick Off 3/25</td>
</tr>
<tr>
<td></td>
<td>• Analyze historical templates to determine existing logic for updating RDC</td>
<td></td>
<td>• High level process mapped</td>
</tr>
<tr>
<td></td>
<td>• Identify data stewards, report owners, and data custodians for process needs</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Document current logic and prepare to pilot first run of process</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation Phase I:</td>
<td>[Determine Current State](Est: 4wks, 03/16 – 04/24)</td>
<td>In Progress; On Schedule</td>
<td>• Communicate RDC changes to ETL team</td>
</tr>
<tr>
<td></td>
<td>• Pilot RDC process</td>
<td></td>
<td>• Document process with Job Aide</td>
</tr>
<tr>
<td></td>
<td>• Identify gaps in RDC</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Partner with data stewards and data custodians to determine placement of</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>codes/names</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementation Phase II:</td>
<td>[Build Future State](Est: 5wks, 04/27 – 05/29)</td>
<td>Not Started; On Schedule</td>
<td>Not started</td>
</tr>
<tr>
<td></td>
<td>• Automation of RDC by NYU IT (Lead by IT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Post-Implementation:</td>
<td>[Pilot &amp; Train](Est: 4.5wks, 06/01 – 06/30)</td>
<td>Not Started; On Schedule</td>
<td>Not started</td>
</tr>
<tr>
<td></td>
<td>• Formalize RDC process</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Create standards, procedures, workflows, etc.</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Ensure risk controls are established for RDC</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Office of Data Governance
• Tammana Sinha, Data Governance Analyst
• Oksana Kinker, Director of Data Governance and External Reporting

Office of Institutional Research and Data Integrity
• David Vintinner, Executive Director of Institutional Research

Office of the University Registrar
• Al Gentile, Associate Registrar

Office of Budget and Financial Planning
• Scott Harlow, Director, Budget & Planning

NYU IT – DRM Support
• Ram Adiga, Senior Systems Analyst, Hyperion / Financial Planning (Technical Data Support of DRM)

UDW+
• Andy Maliszewski, EDM Program Director (Technical Data Support of UDW+)

Reporting
• Addis Crooks-Jones, Director, Information and Analytics Services

NYU IT
Data Governance Support Services
• Adrian Kulpa, Senior Director, Data Governance Support Services and Master Data Management
• Huey-Chih Lee, Project Manager
• Satya Kunta, Program Director, Technology Architecture, Data Management & Reporting
• Kasturi Sen, Data Governance Service Lead, UDW Business Intelligence

Communications & Change Management Office, NYU IT
• Monica McSharry, Senior Portfolio Liaison

WorkDay
• Rajiv Singh Director, Employee Systems Management
• Dhaval Patel Manager, Workday Applications Application Development

FAME
• Raymond Lau - Associate Director, Financial Systems

PeopleSoft Campus Solution
• TBD
Collibra DGC Update
We are making progress toward our goal of retiring the NYU Business Glossary in ServiceLink!

We are giving you an opportunity to review before we communicate the transition to Collibra to the larger NYU data community.

Why?
1. It is a great way to get acclimated with the new tool
2. We may not have captured everything
3. You may see terms that are new to you
4. You may want to fix something out of date

*If you have terms under your domain, we’ll follow up with you directly.*
Business Glossary Interface
1. Default Views / Columns
2. Tags

https://nyu.collibra.com/
#1: Improve trust through transparency

Business terms can be published to the production environment at various stages in the governance process:
#2: Define business terms, not just system attributes

**Official First Name**

*The given name for an individual that can be verified on legal or official documentation.*

**SIS**

*first_name*

**WD**

*Legal name – first name*

**Registry**

*FName*
Announcements & Events
Questions?
Thank you!