Data Governance at Yale

Program Update

October 9, 2018

Last updated: October 2, 2018

Agenda

| Committee Design | Committee Members  
| Data Culture, Goals of Data Governance |
| Responsibilities Review | Committee Role Guide  
| Working Group Role Guide |
| Data Security | Challenges in Securing Integrated Data  
| Review CoRe Security Model Recommendation |
| Next Steps |  |
Data Governance Committee Membership

<table>
<thead>
<tr>
<th>Domain</th>
<th>Steward Name</th>
<th>Role</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Research</td>
<td>Pamela Caudill</td>
<td>Co-Chair/Analytics VP</td>
<td>Tim Pavlis</td>
</tr>
<tr>
<td>Student</td>
<td>Emily Shandley</td>
<td>Co-Chair/CIO</td>
<td>John Barden</td>
</tr>
<tr>
<td>Faculty</td>
<td>Karen Anderson</td>
<td>Data Governance Manager</td>
<td>Maureen Velazquez</td>
</tr>
<tr>
<td>Development</td>
<td>Sondra Cruver</td>
<td>Information Security</td>
<td>Paul Rivers</td>
</tr>
<tr>
<td>HR</td>
<td>Judy Offutt</td>
<td>Institutional Reporting</td>
<td>Ryan Schlagheck</td>
</tr>
<tr>
<td>Finance</td>
<td>Jackie Tucker</td>
<td>OIR</td>
<td>Lily Svensen</td>
</tr>
<tr>
<td></td>
<td></td>
<td>General Counsel</td>
<td>TBD</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Business Operations</td>
<td>TBD</td>
</tr>
</tbody>
</table>

Data Culture – Aspirational State

- Institutional data is a shared community asset; Yale owns its administrative data.
- Data should be accurate, secure and available for appropriate organizational uses.
- Subject matter experts partner with IT to inform and develop integrated data solutions.
- Data quality improvement is institutionally valued.
- Integrated institutional data improves efficiency and provides analytical insights.
- Expert curated data dictionaries support agile analytical self-service.
## Goals

Identify and catalog Yale’s institutional administrative data, stewards, resources and corresponding systems.

Confirm and document institutional field definitions for data and calculated fields. Subject matter experts (SMEs) will be consulted for such confirmation.

Ensure definitions that cross data domains are documented to illustrate the correct cross reference object(s).

Manage access, with appropriate data protections and documented rules to explain any acceptable business exceptions.

Create a focal point for awareness of these definitions and standards (e.g., a website).

Leverage external data models, where appropriate, to support comparative analysis.

---

## Roles & Responsibilities

<table>
<thead>
<tr>
<th>Role</th>
<th>Qualifications</th>
<th>Responsibilities</th>
<th>Examples</th>
</tr>
</thead>
</table>
| Data Domain Steward   | Individuals with decision rights for day-to-day operational and administrative management of source system data, including access rights, scheduling of system updates/upgrade and knowledge of appropriate data use(s) or restrictions. Responsible for data quality remediation activities. Identifies appropriate data SME and resource partners. | Stay abreast of data governance activities (meeting agendas, minutes; items up for approval). Approve domain-specific content developed by working groups. Provide SMEs and resources in support of data governance initiatives. Provide input and votes on institutional policies and procedures. | Research Administration  
Student Space Usage  
Faculty  
Human Resources  
Finance  
Alumni |
| Functional Steward    | Subject matter experts in identified data use function.                          | Participate in data governance meetings when consulted for advice. Contribute advisory guidance on functional area(s) as needed. Review draft content and content up for approval as needed. | Information Security  
General Counsel  
Business Operations  
Institutional Reporting |
| Working Group Members | Understand how data should be interpreted; Able to validate data for accuracy; Able to guide user acceptance testing with test cases; Provide input on data classification and semantic definitions. | Participate in domain data quality activities. Serve as liaison with SMEs as needed to support domain data governance research. Review all domain content for accuracy; seek SME input as needed. Confirm content is developed and approved in compliance with project needs. | OIR  
User Acceptance Testing  
Missing Data Corrections  
De-duplication  
Data Dictionary edits  
Definitions  
Classifications  
Preferred source per item |
| Data Governance Manager | Institutional data asset advocate; Aids data quality remediation; Facilitates data stakeholder forums. Experienced enterprise data steward. | Enterprise communication and coordination. Provide logistical support to each stewardship group. Facilitate development of data governance standards and processes. Review all domain content for compliance with approved standards. Ensure compliance with approved processes. Ensure coordination across data domains. | Data knowledge portal  
Data quality scorecards  
Data knowledge artifacts  
Best practice training guides |
Data Security

• Promote accountable data protection through:
  • Catalog of associated data classifications (with SME confirmation);
  • Publication in the data dictionary for CoRe data assets.

• Data Governance Committee approved security policies and procedures for integrated analytics data and tools.

• Security administration proposal:
  • Domain stewards inform and enable appropriate integrated data use.
  • Stakeholders review and jointly recommend administrative process.
CoRe Security: Approach for RE

Current Source Systems
- Different Transactional Needs
- Different Access Rules

CoRe Security Principle
- Controlled but Uniform Insights

CoRe Security Execution
- For all components of requested and awarded proposal budgets, award details, and non-salary expenditures:
  - Users will have access to the superset of cost centers
- For salary expenditures in Workday:
  - Access to summarized salary expenses will have identifying people information based on the Workday role of Cost Center Payroll Details.

Future Committee Effort

- Review and approve data governance policies
- Discuss and confirm procedures for Common Reporting Platform
- Identify resources and confirm Working Group assignments
- Refine and confirm effective data artifacts and tools
- Communicate progress

Note: 52% of people have access to more cost centers in WD than they do in IRES and 24% of people have access to fewer cost centers in WD than they do in IRES.
Q&A

Appendix
CoRe Program Progress

Program Objective

In instances where operational reporting should be supplemented, CoRe will provide a common architecture and tools to maximize economies of scale in technology to deliver integrated analytical and reporting capabilities. CoRe will not supplant existing systems of record and it will not replace operational reporting from such systems where it meets user needs.

Research Enterprise (RE)

Scope: Integrate proposal and award data from Workday to provide an "end-to-end" view of research data across multiple source systems.

- Vetting against data in progress
- Requirement sign-off received
- Data Mart Security design/Build set up in Dev.
- Completed Data Mart data load for proposal detail. Unit testing completed.
- Proposal Detail technical data validation completed
- Proposal Detail Standard report started.
- Test Env stood up, shakedown in progress.

Student

Scope: Integrate data from 25Live and multiple EMS instances. Use cases for this data include classroom utilization analysis and modeling for courses, exams, make-ups, tutoring, office hours, events, identify meeting patterns for courses, etc.

- Requirements sessions held on 9/24, 10/2.
- Upcoming requirement sessions scheduled for 10/9, 10/15, 10/22, 10/29.
- Anticipated system scope includes 25Live, Event Management System (EMS) and FAMS (Facilities Management)

Faculty

Scope: Support faculty career insights and analysis by integrating data from Workday, Interfolio and the Porter database

- Target start – April 2019

Infrastructure & Security

Scope: Support 25Live, Event Management System (EMS) and FAMS (Facilities Management)

- Currently validating TEST environment configuration; DEV shakeout complete 9/30

Security

- Recommendation for viewing integrated IRES & Workday data: For all components of requested and awarded proposal budgets, award details, and non-salary expenditures, users will have access to the superset of cost centers. For salary expenditures in Workday, access to summarized salary expenses will have identifying people information based on the Workday role of Cost Center Payroll Details.
- Recommendation will be drafted in a memo to the IG Committee

Data Governance

- Developing CoRe data security administration policy (RE).
- Adding business-contributed attributes to CoRe data dictionary
- Committee meets 10/9; lunch and learn for REA set for 10/15.

Comms/Change Management

- Coordination with Change Management underway
- Collecting population information for end users of RE Standard Report suite to determine communication and training strategy
- CoRe Introduction article published 9.30.18
- CoRe Website features planning started.

Risks / Leadership awareness

- Finance dataset build is consuming time from DM lead. Activities assigned to lead will be impacted due to conflicts.
- Competing priorities from RE SMEs will delay data validation activities on the data mart.
- Access issues through non Yale devices and VPN access for non ITS need.

Yellow/Red Status

- Finance – Forecast is currently projected at ~5% over budget. We are monitoring and will identify areas of opportunity in Q4/Qearly Q3.

CoRe Program Timeline

July


September

- Milestones: 6. RE Models reviewed 7. ETLs began finalized 8. ETLs standards finalized 9. ETLs sourcing finalized 10. Infrastructure deployed

October


November


Markets


CoRe Website
RAIL – Rolling Action Item List, Follow Ups