Data Cookbook Implementation Team Final Report

April 2019

Implementation Team

The members of the Data Cookbook Implementation Team (DCIT) are:

- Mary Heather Smith, Office of Institutional Research
- Jennifer Lane, CIS Business Intelligence team
- Jeffrey Fox, Division of Biology and Medicine
- Meghan Hall, Office of the Dean of the Faculty
- Wendi Lewis, CIS Business Intelligence team
- Brook Moles, CIS Business Intelligence team
- Roland Hall, CIS Business Intelligence team

Our Charge

The DCIT was charged with implementing a data dictionary for Brown University using the Data Cookbook, a highly regarded cloud-based data governance solution designed specifically for higher education. Data Cookbook was jointly selected for this purpose by the CIS Business Intelligence team and the Office of Institutional Research.

Our guiding principle has been that data is a shared community resource. In the interests of transparency, and to generate interest from potential users, we gave all members of the Brown community access to view the Data Cookbook. In addition, we provided a link for anyone at Brown to request the ability to add definitions.

An early decision was to organize the Data Cookbook by functional area rather than by data system. Those responsible for defining a term will be the functional users who are most familiar with the business processes around that term. This ensures that definitions are written and vetted by those who best understand their nuances and how those terms are used at Brown.

Outreach to and engagement of the Brown community have been central to our implementation process. In determining which functional areas to start with, we began with Faculty Affairs to capitalize on their interest in this project. We next engaged the Student area since their terms are central to so much of the University and many people in other areas depend on having clear definitions of those terms. To spread the word about the Data Cookbook, we have given status updates and demonstrations to the Institutional Reporting Working Group, an ongoing group with members from across Brown.

What We've Accomplished

We began by breaking down the implementation into sub-tasks to ensure that we would cover all the elements of implementation. As of April 2019, we have successfully gotten the Data Cookbook up and running. Specifically, we have completed the following goals:

Goal	Progress / Status
Define core functional areas, technical systems, and user roles and groups.	**Complete.** Defined 15 functional areas, 9 data systems, and 14 user groups.
Configure workflow to manage approval of both functional and technical definitions.	**Complete.** Configured two workflows—one for a functional definition with no technical definition, and one for a functional definition with technical definition(s).
Enable Shibboleth authentication.	**Complete.** Authentication enabled with auto-creation of new accounts allowed, so any Brown community member may log in and view content.
Establish a location for sharing information with the Data Cookbook user community.	**Complete.** Created Data Cookbook wiki space for sharing meeting minutes, data governance resources, and other information.
Develop a definition style guide.	**Complete.** Style guide shared in Data Cookbook wiki space.
Develop training for definition editors and approvers.	**Complete.** Training handouts and an outline for in-person training have been developed.
Draft and obtain approval for 25 definitions in the Faculty Affairs functional area.	**Complete.** 49 functional definitions approved.
Draft and obtain approval for 25 definitions in the Student functional area.	**Complete.** 41 functional definitions approved, 16 in drafting stage.
Develop integration with Cognos and Tableau content.	**Complete.** We developed a process to insert Data Cookbook definitions into Cognos reports and Tableau dashboards. Documentation of these processes is available in the Data Cookbook wiki.
Develop a definition lookup tool and embed it on the Data Governance website.	**Complete.** Anyone may look up approved definitions on the Data Governance website without logging in.

What Remains to be Done

While the Data Cookbook is fully implemented and being used by the Brown community, several of our initial sub-tasks are not yet completely finished:

Goal	Progress / Status
Engage users in five key functional areas: Admission, Financial Aid, Financial Services, Graduate School, and Advancement. *[Engagement is defined as both editors and moderators having been identified and having logged in, with at least one functional definition drafted and approved.]*	**Work in progress.** Five key areas have been engaged: Admission, Financial Aid, Graduate School, Student, Faculty Affairs. In addition, outreach has been made to two of our original goal areas: Advancement and OVPR.
Develop Cognos reporting capability on Data Cookbook activity and content.	**Work in progress.** A Cognos reporting package is being developed. Will clarify reporting needs with the DCIT. Reports to follow.
Develop reporting and analytics to assist with oversight of content development, e.g., track number of definitions in each stage by functional area.	**Work in progress.** A Tableau dashboard has been drafted and will soon be published and shared with the data stewards.

Next Steps

Now that implementation is complete, the DCIT has dissolved. Ongoing administration of the Data Cookbook will be the responsibility of the Office of Institutional Research, with technical support from the CIS Business Intelligence team. Yet to be determined is the organizational structure for making policy decisions and encouraging or even mandating the ongoing use of the Data Cookbook throughout Brown University.

Recommendation

Ongoing adoption and development of the Data Cookbook will require a sustained effort to educate the community and encourage its use—both by content developers and by content consumers. It will be important to facilitate effective communication among stakeholders in different functional areas in order to ensure that data dictionary terms are appropriately and consistently named and clearly, correctly, and completely defined.

The Office of Institutional Research will lead that effort of facilitating communication between stakeholders and training new users from new functional areas. However, it's important that there be an overarching group with oversight responsibility for data definitions. This group must have the authority to resolve conflicts and set policy for the Data Cookbook going forward.

The DCIT believes that the simplest and most effective means of ensuring the long-term success of Brown's effort to develop a data dictionary is to establish a committee of stakeholders from the functional areas represented in the Data Cookbook.

This group would logically be composed of data stewards and/or their delegates. Service on such a committee aligns with the responsibilities of a data steward as outlined on Brown's Data Governance website:

A data steward is a staff member with oversight responsibility for a subset of the university's data. The steward is typically a functional end user within an operational area who is deemed an expert regarding data managed by that operational area.

Major responsibilities:

- **Implement data standards.**
  Ensure that staff who maintain data are trained to follow standards.
- **Monitor data quality.**
  Work with technical and operational staff to create a process for identifying data entry errors and correcting the data to match university standards. Report to the data trustee any issues that may require larger action on behalf of the university's data governance structure.
- **Handle inquiries about data.**
  Receive and respond to any inquiries related to data that originates from the area they oversee; e.g., questions regarding access, standardization, organization, definition and usage, etc.

If the Data Governance Committee adopts this recommendation, the DCIT would be happy to draft a committee charter for submission to and approval by the Data Governance committee. **We believe it is important that the committee charge officially emanate from the Data Governance Committee; that will convey the importance of the work and increase the chances of its being taken seriously and given high priority.**

Appendix

During the implementation process, we recognized that certain policy decisions are beyond the scope of the DCIT. We have compiled a list of those items, to be decided by the Data Governance Committee or the Data Stewards group mentioned in our recommendations above. Some items on this list are general data governance policy issues, others are more specific questions about how best to use optional features of the Data Cookbook.

1. Should the Data Governance Committee require that project plans for implementing new systems include the Data Cookbook?
   a. Making the definition of terms a part of all new system implementations will lead to better data quality and shared understanding of terms..
2. Should Brown Implement the Data Cookbook Information Request Feature?
   a. The Information Request feature allows any viewer of the Data Cookbook to submit a request for a term to be defined. The Information Request button is on the search page, so if a search for a term is unsuccessful, the viewer could then request the definition.
   b. PROS: having suggestions for definitions come from real users makes it more likely that the definitions are needed and useful. It can be a good way to spur on definition editors to create new definitions.
   c. CONS: if a functional area doesn't respond to a request in a reasonable time, viewers will lose confidence in the Data Cookbook. Some functional areas are not yet active users of the Data Cookbook and may not be read to draft definitions.
3. Should Risk Classifications be Added to All Definitions?
   a. The Data Cookbook includes a feature –Classification Codes—designed to indicate the level of privacy or risk associated with a data element. They are currently an optional feature for editors as they write definitions. Should they be required?
   b. PROS: Good way to familiarize all data users with Brown's risk levels
   c. CONS: Risk and privacy aren't relevant to some data elements (Course Section, Rank, Aid Year). Requiring them to be added to all definitions could be pointless extra work.
4. Should Functional Areas Be Encouraged to Add a Glossary of Data Cookbook Definitions to Their Websites?
   a. We could use Tableau or Cognos to create a list of terms relevant to a given area (Registrar, Financial Aid, etc.). Users could click on the term to see the Data Cookbook definition
   b. PROS: This could allow data consumers to look up definitions for themselves, cutting down on questions to the area and helping to ensure that data elements were being used correctly. Also, as a functional area saw that their definitions were being viewed, they would be motivated to write more definitions.
5. How Can We Improve the Tableau Dashboard of Data Cookbook Use?
   a. We've built a dashboard to track how many definitions exist in different functional areas, how many views each definition has, and where definitions may be stuck in

the workflow. It would be helpful to have feedback from the Data Governance committee or the Data Stewards about this dashboard.

6. Should Brown Standardize Use of School Codes Across Systems?
    a. There are multiple systems of coding schools: CEEB, OPE ID, Title IV Institution Codes, IPEDS ID). Should Brown use a single system for all of our data systems? If so, which system should get precedence?
    b. Do these systems of school codes and the rules for which to use belong in the Data Cookbook?