Data Stewardship Committee: Minutes of February 2, 2015

In attendance: Craig Abbey (OIA), Gary Pacer (EAS), Kelly Hayes-McAlonie (ARC), Sue Krzystofiak (HR), David Love (SEAS), Mark Molnar (OIA), Leah Feroleto (SW), Troy Joseph (GEMS), John Gottardy (Financial Aid), Kim Yousey-Elsener (Student Life), Kelly Kostek (Student Accounts), Jeff Brady (Campus Dining and Shops), Kara Saunders (Registrar), Brian O'Connor, (CAS), Dean Hendrix (Libraries), Ashley Kravitz (Resource Planning), Peter Elkin (Biomedical Informatics), Rachel Link (OIA)

Meeting called to order at 4:01 p.m. by Gary Pacer.

The Data Governance Stewardship Committee, or DSC, is a subsidiary of the Data Governance Council (DGC). Craig Abbey and Gary Pacer are co-chairing the DSC, and the initial invitation sent to the DSC members included the charge from the Provost for both the DSC and DGC. Thirteen data domains are represented on the committee, providing a university-wide perspective of how data governance is managed across the institution, how data are defined for particular projects, and identifying common vernacular to solve or address problems. The DSC membership is comprised of Data Stewards, Ex Officio members, and Ad Lucem (Latin for "to the light") members, who serve to help lead the data stewardship process.

Gary next presented an overview of the process that led to the formation of the DGC, and by extension, the DSC. At the Winter Senior Leadership Retreat held in February 2014, a comment was made that all in attendance agreed upon: a data governance process was needed. Senior leadership agreed that it was hard to discuss strategic leadership decisions if there was no agreement on data definitions or processes. Under the direction of the Chief Information Officer, an executive leadership committee was formed: the Data Governance Executive Oversight Group composed of Vice Provosts and leaders from academic units.

In June 2014, the Guiding Principles for Data Governance (distributed at today's meeting) were approved by the Provost. These Principles provide guidance, but are written at a high level without explaining what tools would be used to access data, or agreement on terminologies that may differ by area (for example, data custodian versus data steward).

In July 2014, an Ad-hoc committee on Roles, Procedures, and Standards for Data Management was formed. This committee identified common terms like data trustee to indicate the Vice-Provost position in charge of data in a particular area, like Laura Hubbard for financial data, or Scott Weber for student records data.

In October and November 2014, the Data Governance Executive Oversight Group, after meetings with the President and Provost, set out the official charge to accomplish definitions of data governance and stewardship by June of 2015.

The co-chairs next provided an overview of the purpose of data governance at UB, which is to support the President, Provost, Deans, University Leadership, and the campus community in informing strategic planning for UB 2020 and beyond; managing daily education and institutional operations through evidence-based decision making; and supporting regulatory, accreditation, governing body, community engagement, and other reporting requirements. There are eight guiding principles for data governance: recognition of primary data sources; unambiguous transformation of data, transparency, data quality, integrity, and security; systemic and continuous; feasible, alignment of data management approaches; and evaluative. The co-chairs stressed that the initial discussion for the DSC is intended to be at a high level to allow focus on the larger picture. Data ownership was discussed next, with the emphasis that no one office "owns" any particular data. Instead, UB is the owner of all institutional data. Individual units may be in trust and responsible for data maintenance, providing space, etc. for the data but UB is the owner of all data.

The initial institutional data focus for the DSC will be on the areas of student records, financial aid, admissions (both undergraduate and graduate), human resources, finance/budget/expenditures, research, space, publications/citations, and organizational structure (entity/hierarchy and academic structure). Examples of operational versus strategic data usages were provided: for example, enrollment can be used to allow access to class rosters (operational) or to calculate faculty workloads (strategic). Examples of data definitions leading to very different numbers were also provided: depending on terminology, calculating student headcounts for Fall 2014 can yield numbers as high as 35,000 for term-activated head count to 28,100 for regular enrollment headcount.

The meeting next moved to discussing the tasks that lie ahead, and how they will be accomplished to meet the June 2015 deadline set forth. The DSC is charged with preparing a report recommending permanent data management oversight and structure which will require:

- Researching what other institutions do;
- Formalizing definitions of data steward/trustee roles;
- Select 3-5 elements from each domain to:
    - Build an inventory of definitions
    - Define attributes like system of record or transformations
    - Store materials in an accessible repository
    - Begin the work of conforming UB institutional data warehouse to official definitions

The committee chairs are considering a website for these materials in the interests of accessibility and transparency.

By April 2015, the DSC must prepare a progress report to the Data Governance Council, prepare an inventory of key data elements, and provide preliminary definitions of students, degrees, and faculty. By June 2015, the recommendations for the permanent data management structure and institutional definitions are due, and implementation will begin.

Working Teams were announced, and Craig and Gary stressed that volunteers were welcome for any team. These teams will help provide the information for the first report to data trustees, due February 26.

Working Team 1: Structure, Roles, and Responsibilities:

- Craig Abbey (Lead)
- Gary Pacer
- Laurie Barnum
- Chris Connor
- Kim Yousey-Elsener

Working Team 2: Enrollment and Degrees (Student Records)

- Kara Saunders (Lead)
- Brian O'Connor
- Greg Olsen
- Mark Molnar
- Lynn DePasquale
- Troy Joseph
- Leah Feroleto

Working Team 3: Faculty (HR):

- Sue Krzystofiak (Lead)
- Tom Okon
- Gary Pacer
- Lauren Young
- Tilman Baumstark

All members will need to assist with the inventory of key institutional data, using a template provided by the committee. The definition of key institutional data is that which UB makes decisions from or is used to monitor institutional performance. These data are the ones that will need to be defined and made available. This process will help with the scope of the project.

Next steps:

1. Implement DSC website
2. Next meeting: February 23 (tentative)
3. First report to data trustees due February 26

Comments from attendees: Kara Saunders asked if the working groups will need to meet prior to the next meeting, and it is affirmed that they should to move forward and show progress. Some information can be done without meeting, like collecting definitions of data elements or starting data dictionaries.

Peter Elkin spoke about definitions, and mentioned that he has learned that while people tend to agree on high-level definitions, we should bear in mind that group membership cans change over time and the numbers of definitions will increase. If the DSC makes a commitment now to correctly define data and represent knowledge, this will pay off in the future. Craig Abbey and Peter Elkin will meet to further discuss this question.

Gary Pacer mentioned that the hope is that enough learning is going on in subgroups that the experience there is useful for other areas and identifies what works and what does not. This will be an evolutionary process.

David Love said that the documentation is ahead of where we as an institution were six months ago. Leah Feroleto mentioned that sometimes defining data also tells you what it isn't. Troy Joseph spoke about the GEMS perspective and how the process in that office can help inform the DSC process.

Jeff Brady mentioned the efforts of Campus Dining and Shops to obtain enrollment data to schedule staffing or services. His office would like to know how many students are enrolled in winter term, for example, to decide on what food vendors should be open. The systems in Campus Dining are self-contained and they make plans based from the data they collect themselves.

Kara Saunders spoke again that the people who contribute to shaping the data – those who enter things like course information – have a large impact on the quality of the data, and that while there is a reliance on those individuals to carefully, accurately, and responsibly enter data, that does not always happen. People enter data thinking of their own, or their unit's needs, and not what others who view that data might need to see. Kara raised the question asking if data stewards should try to catch inaccuracies, which can be problematic when stewards may not know the accuracy of their data. For example, arranged courses may actually be online, or arranged, and it can be difficult to tell which is correct.

Brian O'Connor stated that the crossover of domains is important to remember. Data can be used in multiple areas. Brian mentioned a Tableau report showing student population across campus at various times of the day. This report could be used for class scheduling, space, and even Campus Dining. It is helpful to keep this perspective in mind.

Meeting adjourned at 4:53 p.m.