Information Technology Planning Board Meeting

Meeting Summary | Thursday, February 19, 2015 | 2:00 PM – 4:00 PM | Math Sciences 5628

Attendees:
John Mamer (Chair), Kathleen Bawn (Vice Chair), Apurv Awasthi, Chris Testa, Jason Cong, Jean-François Blanchette, Jim Davis, John Riley, Paul Phillabaum, Peter Kovaric, Robert Trelease

Absentees:

Guests:
Adrienne Dellinger, Christine Borgman, Kelly Wahl

Resources:
Andrew Wissmiller, Kelly Arruda, Kent Wada, Larry Loeher, Ross Bollens

1. Approval of October 27, 2014 Meeting Summary | John Mamer
The Summary from the October 27, 2014 meeting was approved.

2. A UCLA Academic Data Landscape | Christine Borgman, Kent Wada, Jim Davis, Kelly Wahl [Action: Information and Discussion]

Data: An Institutional Asset and Risk
As an institution in the data and knowledge business, the University regularly manages data in the areas of education, research, decision-making, and various services. While it has been traditionally managed as a risk, the new age of data proliferation encourages us to also view the tremendous amounts of data we generate as an asset.

We have been steadily working at a detailed analysis of data itself and its relationship to privacy and security. Privacy and security, with natural tension between them, have been purposely separated as distinct entities in order to assist in balancing. The work of the UC Privacy and Information Security Initiative developed a definition model, which distinguishes between autonomy and information privacy, as well as, information security and security of infrastructure.

KPMG’s higher education audit committee has observed a collision between securing sensitive data and the use of ubiquitous open source and/or low cost applications, tools, and social media. And its becoming increasingly difficult for individuals to understand what their rights and responsibilities are regarding use, storage, and sharing of data when moving from application to application throughout any given day.

The University sits in a network-mediated era of data, which according to recent research by Professor Christine Borgman and others, is about one generation behind the current cyber infrastructure-mediated data space\(^1\). And as is expected, our policies are even slower to adapt to technology changes.

There are several data categories within the institution that stand out for generating tremendous amounts of valuable data: medical, faculty, and student data. And there are common themes among

\(^1\) The report can be found at: http://www.nsf.gov/pubs/2008/nsf08204/nsf08204.pdf
them: the business of open architecture market, valuation of data (collective vs. proprietary), and converting individual data into collective wisdom, innovation, and practice.

**Medical Enterprise and Sustainability:**
1. UC ReX is a systemwide project in which all 5 medical centers are combining patient data for a research exchange; UCLA is the same size as all the other medical centers combined. With proper protections of infrastructure, devices, and data, this project produces extensive analytics that can lead to meaningful development and improvement of health deliveries, financial and health advantages, new discoveries, and innovative practices.
2. NANTHealth Cloud Structure and Acquisitions is an organization that collects an extensive amount of health data with the capacity to convert it from data point to information, knowledge, and eventual wisdom.

**Educational Enterprise and Sustainability:**
1. UC Innovative Learning Technology Initiative (ILTI) is a project aimed at developing cross enrollment of courses across the system. At UCLA, CCLE is a communication hub that already generates vast amounts of data, including course enrollments, content, grades, etc. How can we customize experiences based on these data points, and to what extent can data be shared beyond campus?
2. Unizin is a data depository for a collective of 5 campuses across country.

**Faculty Enterprise and Sustainability:**
1. Opus, currently in development, is the UCLA faculty information system that will aggregate faculty data from across campus. In order to create Opus, a book of records process was developed to reconcile data and ensure accuracy and consistency.
2. Publisher, Elsevier, takes published data, mines and analyzes data points, and monetizes the analysis by selling it back to educational institutions. In essence, we are buying back data about ourselves. What are the metrics used in this data extraction, and do we want it decided for us by the publishers? How do we get involved in the area of developing these metrics for ourselves?

These projects lead to a number of questions about the value of our data in the commercial space, and what sort of partnerships we should be engaging in. It raises questions about ownership, access, privacy, policy, and obligations to public records requests. As we shape the opportunities and risks involved in data valuation, we will need to engage in a thoughtful, multidimensional balancing process, and reflect ethics, scholarship, and institutional values into the contracts, licensing, and partnerships that we establish.

**Data Management, Governance, and Policy Issues**
There are particular concerns around research and publication data. Some research data issues include data management, curation, and stewardship, as well as, human subjects regulations and open records laws. Meanwhile, publication data concerns involve library management, ownership, credit and attribution, use and reuse, public-private partnerships, etc.

For instance, open access to data, such as in eScholarship, is inherently more democratic and equitable, but it also becomes an open field in which business models are built. As businesses employ varying metrics to the open data, the results (e.g., of publication and citation numbers, etc.) vary widely by source. This, of course, raises serious concerns about accuracy.

These result in significant data management, governance, and policy issues for the university. The
depth of these issues in an ever-growing data environment, have led to the establishment of the Data Governance Task Force. The group is investigating how we should collect, organize, and use research analytics about ourselves; who should have access to the data; and what should data governance principles and processes look like.

**Institutional Research at UCLA**
Institutional Research, part of Academic Planning and Budget, is responsible for the systematic collection and analysis of data for accountability reporting (system and state), academic and operational planning, and peer benchmarking. Institution Research is a distributed function across campus and requires IRB exemption.

For example, the research of undergraduate admissions can provide information on student success/standing, program fitness, resource needs, mandatory reporting requirements, and even some components of US News rankings.

Institutional Research collects data from query results, survey data, tabular reporting, test statistics, etc., and they have created a dashboard tool for using some of this data.

**The “Should we?” Questions**
While the work of Institutional Research requires an IRB exemption, how does the institution address stewardship of data functions that fall outside the purview of the IRB (i.e., the just because we can doesn’t mean we should items)? Will the lack of clear guidance signal an automatic green light for those wanting to move forward with use of data?

These ethical challenges are part of what is being considered by the Data Governance Task Force.

‘While big data can be used for great social good, it can also be used in ways that penetrate social harm and render outcomes that have inequitable impacts, even when discrimination is not intended. Small biases have the potential to become cumulative, affecting a wide range of outcomes for certain disadvantaged groups. Society must take steps to guard against these potential harms by ensuring power is appropriately balanced between individuals and institutions.’

Big Data: Seizing Opportunities, Preserving Values | Executive Office of the President (May 2014)

**Moving Forward**
Our distributed nature, along with supporting academic freedom and the obligation of transparency as a public institution, make untangling the data landscape at UCLA extraordinarily complex. While we have been doing good work – UCLA is the first campus to have an evenly distributed joint faculty/staff IT Planning Board and Board on Privacy and Data Protection (and others, see slide #20 of presentation: 01-UCLA Academic Data Landscape-Davis) – there is still a lot of work to do.

**Projecting Values:**
We esteem our faculty, staff, and students individually and as a community; we support a robust culture and set of ethics; and we appreciate the intellectual and financial value of our data.

While we make data available as a public institution, private sources are building business models on that data. How do we wrap our values around our data such that it continues to meet our mission, remains available for making academic and operational decisions, as well as, leverages profitability in an appropriate manner?

**IT governance Role:**
What roles and/or impact can/should IT governance support in the coming years to ensure
appropriate use of data and balancing of security, privacy, openness, confidentiality, policy and management practices that reflect our values?

Let’s take the following suggestions, along with the outcomes of the Data Governance Task Force, to produce more concrete use cases for ITPB to review and discuss.

**Brainstorm of Ideas for Moving Forward:**
Can we pull together some use case to get individuals engaged on particular topics?

- The UC Electronic Communications Policy is up for review this year. When it was first written, email was the foreseeable concern; perhaps the ECP review is an opportunity to infuse some of the data values we have been discussing.
- Create a four quadrant model in an effort to consider multiple factors, such as internal/external resources, better data collection and mining, risk minimization, etc.
- Create an itemized list of challenges and potential resources for addressing them.
- Develop a survey in order to narrow focus.
- Add principles, standards, etc. to review processes, such as CSG’s project review process.
- Bubble-wrap data in values before releasing it into the world, such as with a Creative Commons-like mechanism.