

**Report on the  
Institutional Data Management  
Governance Survey**

**Prepared by Marc Goulden**

**May 6, 2009**

# **IDMG Assessment Report**

## **Executive Summary**

### **Introduction—Survey and Report**

UC Berkeley's Institutional Data Management and Governance (IDMG) Initiative is the campus's step forward in making institutional data—including data related to applicants, students, faculty, staff, alumni, and donor prospects—easily accessible, reliable, consistent, and secure. To spearhead the initiative, George Breslauer, Executive Vice Chancellor and Provost, and Nathan Brostrom, Vice Chancellor—Administration, convened a small Institutional Data Task Force composed of a subset of the Chancellor's Cabinet to examine the current situation, best practices, and proposed solutions, and then make recommendations. In 2008, the Advisory Group appointed to assist the Task Force in this effort conducted the *Institutional Data Management and Governance (IDMG) Survey* to provide a deeper, detailed understanding of the University's overall data environment, and thereby serve as a key foundation for the Institutional Data Task Force's final recommendations.

The *Survey* ambitiously addressed many aspects of the campus's data environment. Conceived as a vehicle for making the *Survey*'s findings usable for the campus community, this *Institutional Data Management and Governance (IDMG) Report, 2009* presents and analyzes the findings of the *Survey* as well as derives from them key preliminary recommendations for future improvements to the campus's "data process." For practical reasons, the report excludes an in-depth discussion of data governance issues on the campus, an assessment of the financial cost of any of the proposed recommendations, an in-depth analysis of open-ended responses to the *Survey*, and a specific set of next steps (as these cannot be established without the participation of other stakeholders).

### **Respondents and Response Rates**

Overall, 394 individuals—representing a broad array of job positions and functional roles relative to institutional data, from members of the Chancellor's cabinet to data recorders—responded to the survey over a three-week period. A number of these individuals expressed appreciation for the opportunity to respond to the *Survey* and noted the importance of the issue to the campus.

In general, respondents directly associated with the *Survey* and the campus responded at a 50% or higher rate, and chief administrative officers responded at a relatively high rate (55%). In contrast, members of systems-related groups responded at lower rates. Respondents from administrative units are

somewhat overrepresented, whereas respondents from academic units/departments are somewhat underrepresented.

### **Assessment of Campus Data Environment—General**

Respondents were asked to rate the data environment with regard to 22 different attributes divided into four broad categories: 1) availability/accessibility of data, 2) usability of data, 3) analysis/reporting of data, and 4) protection of data and privacy. Key results are presented below.

Taken as a whole, Berkeley's campus data environment received discouragingly low ratings. "Necessary computer safeguards/systems to protect against data theft" is the only item where more than 70% of respondents indicated that they believed Berkeley's situation was excellent or good (71%), with only 14% of the respondents evaluating this situation as excellent. On a majority of the items, 12 out of 22, respondents were more likely to evaluate the current situation as fair or poor than they were to rate it as excellent or good.

The items that were evaluated most favorably, receiving at least a 60% excellent/good rating, fell into two broad categories: 1) security of data, and 2) existence/accuracy of data.

The items that received middle rankings (45–55% excellent or good) relate to issues of accessibility of data: having up-to-date data when it is needed, and having the necessary campuswide procedures to support the sharing of data across units. Although these items earned middle rankings, the aggregate campus ratings on them is quite low, with almost half of the respondents evaluating all of these items as fair or poor.

The five items that received the lowest aggregate rating, with more than two-thirds of the respondents evaluating the situation as fair or poor, were concentrated around the analytical portion of data work, including issues related to reporting tools, access to specialized reports, clear documentation regarding data, and consistency of data fields across the campus.

### **Assessment of Campus Data Environment—Specific**

#### *By Major Areas of Decision Making*

Respondents who make or support decisions in the **undergraduate and financial areas report the highest use of data systems**—local, campus, or systemwide—whereas those working in the **research and other population areas report the lowest use**.

Individuals working on financial issues display either the highest or second highest overall rating for each item evaluated in the *Survey*. Individuals working on course-related data seem particularly distressed by the current Berkeley data environment.

Clear procedures for requesting data access appear to be of particular concern across a number of the major decision-making areas, including the following major areas: undergraduate, graduate student, faculty, academic staff, and research.

A good number of respondents are not using data from campus or local systems in support of the decision-making process.

#### *By Sub-Topical Areas of Decision Making*

In general, the **existence of data is not a major bottleneck** on the campus in terms of supporting informed decision making. Rather, **the subsequent portions of data work appear to be of greater concern**, including gaining access to data, understanding the meaning of it, conducting methodologically sound analysis, and converting it to meaningful information that can inform decision making.

Centralized campus systems, particularly ones related to financial data, are associated with more-favorable ratings with regard to established procedures for requesting access to data. Course-related systems and some undergraduate systems and sub-topical areas are associated with lower ratings.

Gaining access to necessary data appears to be particularly problematic for staff who support departmental chairs.

Although the integration of financial data across the campus appears to be associated with more-favorable ratings with regard to release of up-to-date data, topical areas of decision making that are complex and either human-centric (mentoring, climate, productivity, etc.) or require longer-term planning (e.g., staff succession planning, hiring policies, proposal trends, etc.) are associated with more-negative ratings.

Systems managers and policy analysts are particularly likely to rate consistency of data fields across systems in the negative, suggesting that this is a particular area of concern for the campus (given their significant expertise in this area).

Access to user-friendly reporting tools is, overall, rated toward the negative end of the spectrum, and the few bright spots pertain to the budget and the BIBS and BIS systems.

A small number of sub-topical areas are associated with particularly unfavorable ratings on access to analytical tools to help with data, but a couple of high points jump out: OSR's survey systems and research databases.

Institutional research and policy analysts rate existence of necessary data, access to user-friendly reporting tools, and access to analytical tools to help with data more favorably than many other groups.

At present, the campus is rated more favorably for securing data than providing access to data.

## **Looking Ahead—Creating a Better Data Environment**

### *Context: The High Dispersal of Campus Data Needs and Great Variation in Functional Roles*

There is a high dispersal of data-dependent decision-making activities on the Berkeley campus, as well as great variation in functional roles among campus individuals directly involved with data gathering, recording, management, analysis, and use. Success in improving Berkeley's data environment will require meeting the often differing needs and priorities of all involved functional groups, spanning the complete range of decision making and support functions.

Deans and chairs represent one extreme among campus data consumers, as they perform the widest array of decision-making activities.

### *Key Task: Mapping Common Areas of Data System Access/Use*

Based on an analysis of the general patterns of use of existing campus systems, core business functions, student-related systems, and human resource systems hold the most promise for improvement through larger-scale integration.

### *Future Opportunities: Range of Possible Solutions*

Based on our analysis of data from the *Survey*, the campus is hampered by its lack of data integration. Currently, we are neither efficiently nor effectively able to support many campus decision-making areas with useful data-derived information.

Given our current weaknesses and fiscal climate, however, it is not reasonable to expect that we could seamlessly create a new one-size-fits-all solution to address our current inadequacies. **A complementary, multi-level approach appears necessary**, with new initiatives carefully chosen based on their likely campus impact and likelihood of success.

### *Report Summary: Recommendations and Planning Questions*

A rich set of recommendations evolved out of the data findings. Key recommendations include:

- Seek additional input from and actively consult with academic units/departments, deans and chairs, and various functional and data consumer types in future and ongoing data efforts
- Continue to map and analyze the campus data environment, investigating why data is not used in supporting some decisions and considering data-use patterns in future efforts to improve the data environment
- Port successful procedures to problem areas and disseminate inventories of data and analytical resources that seamlessly interface with preferred data tools (such as Excel)
- Investigate any perceptions of data inaccuracy and solve or mitigate any identified problem
- Prioritize cost-effective projects that increase data efficiency and consistency, carefully evaluating projects for their likely benefits and costs (planning questions/metrics are provided to assist with this effort)
- Encourage appropriate integration of data resources and tools to meet local and campus needs

Out of the complete set of recommendations, the Advisory Group designated three overarching ones as most important:

- 1. Make consistency of data fields across systems a priority**
- 2. Address accessibility issues at the department and school/college level**
- 3. Encourage appropriate integration of data resources and tools to meet local and campus needs**

### **A Collaborative Approach for Change**

The Advisory Group hopes that the information and recommendations contained in this report will help the campus collaboratively formulate and implement appropriate improvements to Berkeley's data environment according to a timeline that is both aggressive and practical.