IDMG Assessment Report
The Survey

Introduction

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 (for a full list of Task Force, Advisory Group Members and other functional members of the effort please see *Figure A-1* in the *Appendix* or go to http://administration.berkeley.edu/idmg/people.htm).

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).

The initial plan of the Advisory Group was to hold a series of focus groups, roughly 10 to 15, with key UC Berkeley stakeholder groups to determine the current state of data issues on the campus and future needs (see the potential focus group list in the *Appendix, Figure A-2*). Upon further reflection, the Advisory Group determined that a survey instrument would 1) reach a broader array of campus constituency groups, including many individuals who would not be asked to participate in focus group meetings; and 2) give individuals a greater opportunity to provide useful feedback. Since decision-making activities; data needs, access, and support; and employees' functional roles are highly

individualized, a survey with the flexibility to accommodate highly disparate respondent profiles was identified as the most promising assessment approach.

Design of the Survey

From late November 2007 to early March 2008, the IDMG Advisory Group worked together in the development of the instrument and the identification of the survey population. The final survey was organized into four major sections: 1) Respondent Profile—campus role and use of data; 2) Assessment of Campuswide Data Systems—experiences with UC Berkeley's institutional data systems; 3) Assessment of Campus Data/Research Analysis Environment—experiences with data–decision making/governance and implementation/management; and 4) Final Questions—suggestions for improvement.

As a result of the survey design process, the Advisory Group identified a number of essential and challenging tasks necessary for assessing our campuswide data environment, including clearly defining decision-making areas, functional job titles, and divisional categories; creating a comprehensive and carefully organized inventory of campuswide data systems (fortunately, this task was already well underway thanks to the hard work of the Data Stewardship Council; see http://datasteward.berkeley.edu/); and fully mapping the process whereby raw data is taken and developed into meaningful information to inform campus decision making. After the completion of these tasks, the survey underwent extensive pilot testing and multiple rounds of revisions. The Advisory Group launched the instrument on March 7, 2008.

In the e-mail inviting individuals to participate in the survey, respondents were informed that they should allow at least 60 minutes to record their responses. This liberal estimate reflected the fact that, depending on a respondent's profile, the survey might be quite involved. For individuals with a large range of decision-making responsibilities, or individuals who directly access a large number of campuswide datasets, the number of questions was considerably larger than for other individuals. Employing JavaScript and the impressive technologies developed by Office of Student Research (OSR—see https://osr2.berkeley.edu/), the survey was designed to "drill down" to further questions, as warranted. Respondents were informed regarding this aspect of the survey design by the following statement: "Depending on the responses you select, additional questions will open. If you work with a wide range of institutional data, e.g., as a research analyst, you can expect to have many additional sections open up and we hope you will take the time needed to fully complete the survey." A full copy

of the survey can be accessed at http://gradresearch.berkeley.edu/IDMG survey.html. (Note: To see hidden sections, specific responses must be selected.)

Respondents' Evaluation of Survey

The median time to complete the survey proved to be 32 minutes; third-party anecdotal comments by respondents after-the-fact suggested that it was easier to respond to than some had expected. **Although a few respondents expressed frustration with the length of the instrument or felt it was poorly designed, a larger number of respondents expressed appreciation for the opportunity to respond to the survey.** Representative responses include:

> *Thank you for this survey. It's the first time that anyone has inquired about my data needs.*
>
> *I found this survey to be somewhat difficult to plow through.*
> *I know how hard it was to design this questionnaire since it is covering so many different systems. Good first effort!*
>
> *I experienced survey fatigue.*
>
> *I am glad to see that we are finally working on this important issue. Please count me as a big supporter of this initiative. Let's move forward!*
>
> *Thanks for asking, this survey is a great vehicle.*

In an effort to assure frank responses to the inquiry, respondents were pledged confidentiality. Even so, some individuals may have worried that their verbatim comments might eventually be attributed to them and may have been reluctant to refer to specific individuals or units when discussing their experiences, particularly negative experiences. In fact, a review of open-ended comments in the survey suggests a relative dearth of negative comments directed at specific individuals or units—relative to informal, off-the-record discussions—suggesting some respondents may have partially censored their written comments (in some cases, confidential follow-up interviews may be required to assure full disclosure of politically sensitive information). Reflecting this general concern about confidentiality issues, one respondent wrote at the end of the survey, "I have been very blunt and straightforward in this survey and trust that my comments will be kept anonymous."

Respondents

Thumbnail: The survey benefited from a large and diverse respondent pool.

Introduction

The initial sample population was identified based on membership lists of the relevant groups (see the Potential Focus Group list in the *Appendix, Figure A-3*). Upon receipt of the invitation e-mail, this initial group was asked also to forward the e-mail to others they believed should respond to the survey. All respondents logged-in via CalNet authentication so that we knew who had responded to the survey and so that individuals who were not originally in the survey population could be allowed access (because of the confidentiality pledge, the link between respondent information and actual responses was automatically detached). This use of snowball sampling—initial respondents recruiting additional respondents from their work associates—limited our ability to calculate overall response rates. It did, however, assure a broader range of respondent types and a larger number of respondents than might have occurred had a more conventional approach been adopted.

Respondents received an initial invitation and one subsequent reminder, though the forwarding of e-mails to relevant parties resulted in some receiving multiple e-mail invitations, and then multiple subsequent e-mail reminders—even if they had already responded to the survey. Overall, 394 individuals responded to the survey over a three-week period.

> ***Observation: High Dispersal of Data Use and Needs.*** *The large number of respondents [to the report] suggested to some members of the Advisory Group that data use and needs on the campus may be more dispersed than they initially anticipated.*

Profiles

Figure 1 shows the number of respondents by their *position relative to institutional data* on the Berkeley campus. Because it was determined early in the design of the instrument that some respondents were likely to wear multiple hats with regard to data issues, respondents were asked to declare all roles that applied to their situation (thus a person may appear more than once in *Fig. 1*).

As can be seen by *Figure 1*, **survey respondents spanned a broad array of job positions relative to institutional data, from members of the Chancellor's cabinet to data recorders.** The total number of roles cited by respondents was 728, with respondents averaging 1.8 roles. Detailed

analysis indicated that 60% of respondents stated that they had only one role in regard to institutional data, another 28% had two or three roles, and the remaining 12% had four or more roles. The most likely groupings of multiple roles included (based on factor analysis—principal component analysis with varimax rotation): 1) staff member who supports non-academic departmental decision maker, general analyst, and data recorder; 2) member of Chancellor's Cabinet, and campus-level decision maker; 3) manager of institutional research unit/office, institutional researcher/analyst, and policy analyst; 4) college/school-level leader (e.g., dean), and academic department leader (e.g., chair); 5) systems manager, and systems programmer; and 6) college/school-level administrator (e.g., associate dean), and academic department administrator (e.g., associate/vice chair).

Figure 1: Position relative to institutional data
Which of the following describe your position relative to institutional data?
(check all that apply)
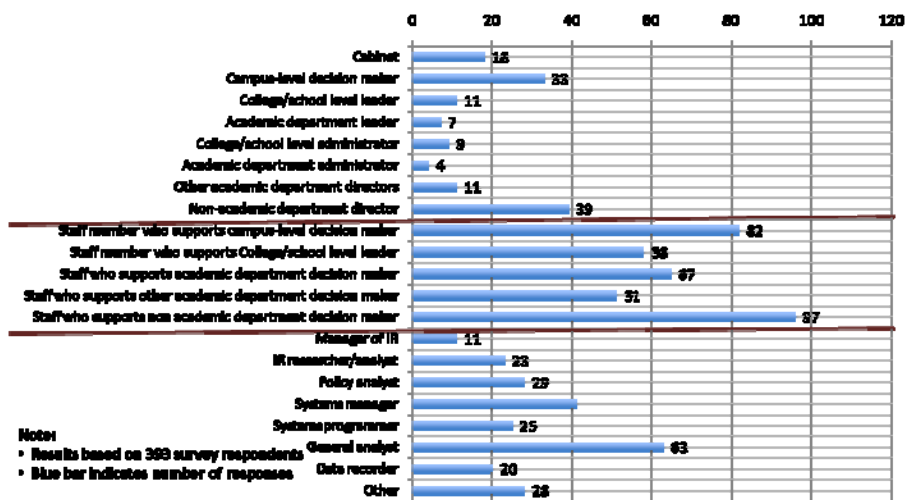
Response Rates

By Campus Group

The response rates for known relevant campus groups varied widely (see *Figure 2*). In general, groups directly associated with the project (i.e., the IDMG Taskforce and Advisory Group) and the campus leadership, including the Chancellor's Cabinet and Council of Deans, responded at a 50% or higher rate. So, too, chief administrative officers (the CAOs or MSOs of campus units/departments) responded at a relatively high rate (55%). All of these groups were directly encouraged through various

means to respond to the survey (presentations at group meetings, e-mails, and personal invitations). Members of systems-related groups, including the various Student Systems groups, the Campus Information Security and Privacy Committee, and the Information Technology Architecture Committee, responded at lower rates.

Figure 2: IDMG Response Rates for Known Relevant Campus Groups

Committee/Organization	# of Respondents	# Surveyed	Response Rate
IDMG Task Force and Advisory Members	22	22	100%
Student Systems 2012 (Exec. Governance Member.)	7	7	100%
Chancellor's Cabinet	19	22	86%
CC2	11	17	65%
Student Systems 2012 (Project Team)	3	5	60%
Chief Administrative Officers	30	55	55%
Council of Deans	12	23	52%
Data Stewardship Council	13	25	52%
Undergraduate Deans	6	12	50%
IR/Managers & Staff	4	8	50%
Student Systems 2012 (Community Council)	13	28	46%
Campus Information Security and Privacy Comm. (CISPC)	8	21	38%
Associate CIOs	3	8	38%
Student Systems 2012 (Technical Subcommittee)	4	11	36%
Student Systems 2012 (Vision Working Group)	2	6	33%
Student Systems 2012 (Financial Metrics Group)	2	7	29%
Information Technology Architecture Commit. (ITAC)	5	21	24%

By Budget Control Unit

Figures 3a and 3b below show that respondents to the survey come from a wide variety of different budget control units. In fact, there are one or more respondents from every control unit on the campus.

Figure 3a: Budget Control Unit

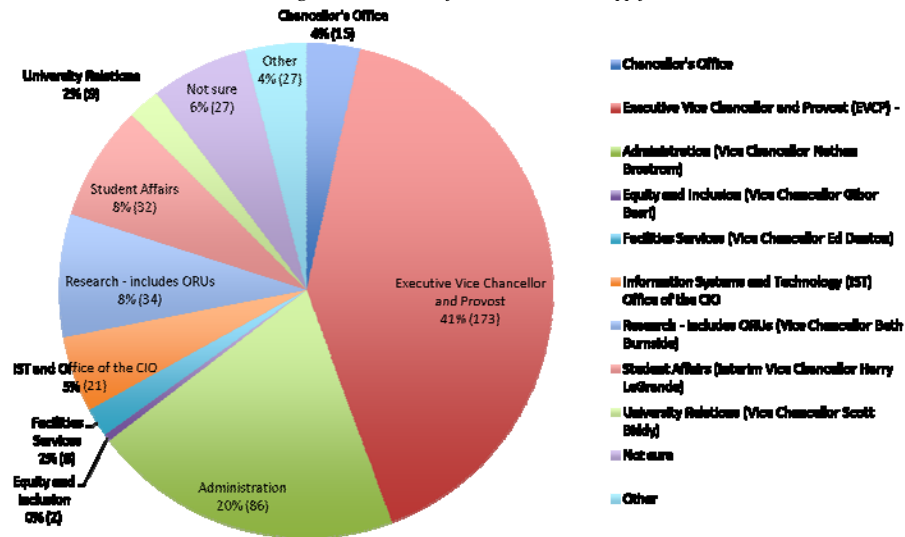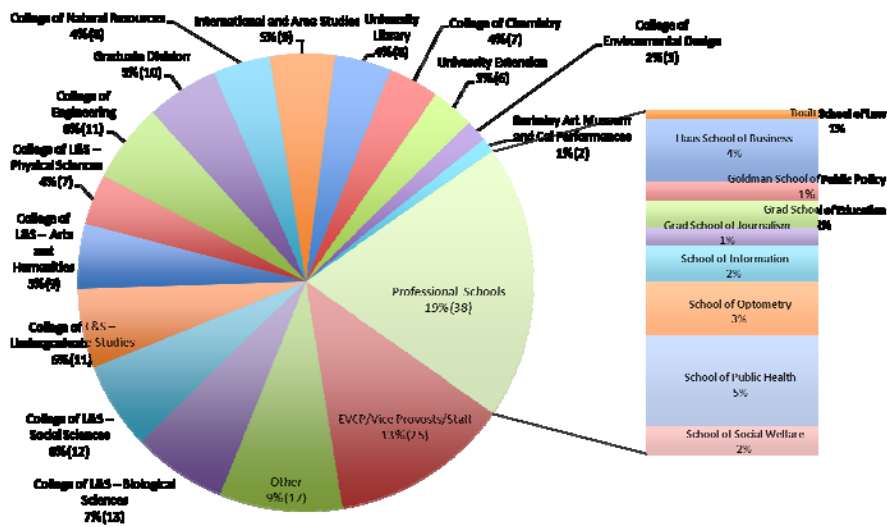*Which budget control units are you in? (check all that apply)*

Chancellor's Office 4% (15)
Other 4% (27)
Not sure 6% (27)
University Relations 2% (9)
Student Affairs 8% (32)
Research - includes ORUs 8% (34)
IST and Office of the CIO 5% (21)
Facilities Services 2% (8)
Equity and Inclusion 0% (2)
Executive Vice Chancellor and Provost 41% (173)
Administration 20% (86)

Legend:
- Chancellor's Office
- Executive Vice Chancellor and Provost (EVCP) -
- Administration (Vice Chancellor Nathan Brostrom)
- Equity and Inclusion (Vice Chancellor Gibor Basri)
- Facilities Services (Vice Chancellor Ed Denton)
- Information Systems and Technology (IST) Office of the CIO
- Research - includes ORUs (Vice Chancellor Beth Burnside)
- Student Affairs (Interim Vice Chancellor Harry LeGrande)
- University Relations (Vice Chancellor Scott Biddy)
- Not sure
- Other

Figure 3b: EVCP Budget Control Unit

*Within EVCP Budget control unit, which departments/divisions are you in?
(check all that apply)*

College of Natural Resources 4% (8)
International and Area Studies 5% (9)
University Library 4% (8)
College of Chemistry 4% (7)
College of Environmental Design 2% (3)
Graduate Division 5% (10)
University Extension 3% (6)
College of Engineering 6% (11)
Berkeley Art Museum and Cal Performances 1% (2)
College of L&S – Physical Sciences 4% (7)
College of L&S – Arts and Humanities 5% (9)
College of L&S – Undergraduate Studies 6% (11)
College of L&S – Social Science 6% (12)
College of L&S – Biological Sciences 7% (13)
Other 9% (17)
EVCP/Vice Provosts/Staff 13% (25)
Professional Schools 19% (38)

Professional Schools breakdown:
- Boalt School of Law 1%
- Haas School of Business 4%
- Goldman School of Public Policy 1%
- Grad School of Education 3%
- Grad School of Journalism 1%
- School of Information 2%
- School of Optometry 3%
- School of Public Health 5%
- School of Social Welfare 2%

Observation: Lower Response Rates among Academic Units/Departments. When compared to the general campus control unit headcount, it appears that the respondents from administrative units are somewhat overrepresented among our respondent population, whereas respondents from academic units/departments are somewhat underrepresented.

Recommendation: Seek Additional Input from Academic Units/Departments. The campus data community should interface more intensively with academic units/ departments to ascertain their data use patterns and needs.

Further Comments

Area of Concern — Department Chairs. Only seven chairs responded to the survey out of the 60 or more on campus. (Note: it is unclear how many were sent e-mail invitations). Accordingly, a follow-up inquiry directed at department chairs is likely desirable, given their key role on the campus.

Observation: Diversity of Functional Roles. The large group of respondents was characterized by a great diversity of functional roles (with many individuals playing multiple roles) in regard to data and decision making.

Recommendation: Consult Full Range of Data Consumer and Producers in Future Efforts. Future data management/governance efforts need to pay careful attention to the range of potential data consumers and data producers. Vetting future efforts in light of data from this survey can be a first step in seeking to build solutions that meet the needs of the larger campus community. In many cases, additional in-depth analysis will be necessary.