

## IDMG Assessment Report

### Assessment of Campus Data Environment — General

*Observation: Low Campus Ratings. Taken as a whole, our campus data environment received discouragingly low ratings.*

As part of the survey, respondents were asked to rate, from excellent to poor, the campus institutional data environment based on their “experiences and perceptions” 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. *Figures 4a, 4b, and 4c* display the top eight highest-ranked, middle 6, and 8 lowest-ranked attributes based on respondents’ evaluation of particular items (“not sure” and “not applicable” responses were excluded from this analysis).

The results are sobering. The highest-rated item, “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%); 29% thought the situation was either fair or poor. There is certainly room to debate the level to which Berkeley should aspire in terms respondents’ ratings (for example, the advisory group felt that items with rating of 70% excellent or good were an acceptable threshold based on prior survey research), but the fact that our current most-favorably evaluated item is “computer safeguards” is a less than impressive achievement—particularly given the fact that only 14% of the respondents evaluated this particular situation as excellent. In fact, **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.**

#### *Most-Favorable Ratings: Security of Data, and Existence/Accuracy of Data*

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.

**Figure 4a: Top 8 highest ranking attributes of campus institutional data/environment**

Q#	Attribute	% Exc./ Good*	Total N
C1A-4	Necessary computer safeguards/systems to protect against data theft	71% (14%)	222
C1B-2	Necessary computer safeguards/systems to protect against data loss	68% (15%)	209
	Accuracy/quality of the data you access	68% (12%)	263
C1A-2			
C1A-3	Clearly established procedures to ensure the protection of data	66% (15%)	244
C1B-1	Clearly established procedures to ensure that appropriate employees have the correct access to data	64% (12%)	233
C1A-6	Clearly established procedures that ensure the confidentiality/privacy of individuals	63% (15%)	237
C1A-1	Existence of the data you need to support informed decision making	63% (9%)	301
	Willingness of other units to supply/grant access to the data you need	56% (10%)	272
C1A-5			

**•Note:**  
 "Not sure" and "not applicable" have been excluded from analysis. % indicates respondents who evaluated the attribute excellent or good. Number in parentheses is % who scored the attribute excellent.

The relatively higher ratings with regard to security issues may in part be due to the concerted effort that the campus undertook in the aftermath of the well-publicized theft of a notebook from Graduate Division in March 2005, and the related loss of personal data of 98,000 current and former Berkeley graduate students (for example, see <http://www.cnn.com/2005/TECH/03/29/stolen.laptop/index.html>). In the immediate aftermath of this event, the Chancellor pledged that UC Berkeley would achieve excellence in the arena of data security ([http://www.berkeley.edu/news/media/releases/2005/04/05\\_rjbmemo.shtml](http://www.berkeley.edu/news/media/releases/2005/04/05_rjbmemo.shtml)). Although it is not possible to tell based on our time-specific survey data, this event and subsequent campuswide efforts suggest that **concerted efforts in other data-related arenas besides that of security issues may be possible, and may offer the hope of an improved data environment and more-positive evaluations in the future.**

*Middle Ratings:*

The six items that received middle rankings (45-55% excellent or good) are shown in *Figure 4b*. All six of these items relate to issues of accessibility of data: having up-to-date data when it is needed, either through standardized reports or systems that allow easy extraction of data; and having the necessary campus wide procedures to support the sharing of data across units. Although these items rate better than those on *Figure 4c*, the aggregate campus ratings on these items is quite low, with almost half the respondents evaluating all of these items as fair or poor.

**Figure 4b: Middle 6 ranking attributes of campus institutional data/environment**

Q#	Attribute	% Exc./Good*	Total N
C1B-4	Availability of data when you need it (e.g., weekend or after-hours access)	53% (12%)	240
C1B-3	Release of up-to-date data you need (e.g., availability of most recent headcount/FTE/fiscal data)	52% (10%)	266
C1C-5	Clearly established procedures that encourage the use/publication/sharing of data while maintaining privacy protections	49% (7%)	204
C1B-5	Timeliness of cross-unit response to providing access to data	48% (6%)	252
C1C-1	Ease of extracting/accessing data using preferred methods	45% (6%)	265
C1C-3	Access to standardized reports (written, web-based, etc.) that minimize the need to directly access and analyze data	45% (6%)	248

•**Note:**  
 “Not sure” and “not applicable” have been excluded from analysis. % indicates respondents who evaluated the attribute excellent or good. Number in parentheses is % who scored the attribute excellent.

*Least-Favorable Ratings: Analysis-Related Work*

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. Here, the atomic parts of data are analyzed, sometimes restructured, and synthesized into meaningful information that can be used effectively to inform decision making. Of note, consistency of data fields across systems was the most likely of all items to be evaluated as poor by respondents—with a full 46% rating the situation as poor.

**Figure 4c: Bottom 8 ranking attributes of campus institutional data/environment**

Q#	Attribute	% Exc./Good*	Total N
C1C-4	Access to trained staff to assist in the analysis of data you need	43% (8%)	244
C1D-1	Consistency of data you access across different campus systems/units	40% (4%)	212
C1C-2	Clearly established procedures to allow you to request access to data you need	39% (4%)	284
C1D-4	Access to specialized reports that support complex decision making processes	32% (5%)	217
C1D-3	Access to user-friendly reporting tools (web-based or otherwise) to support the analysis of data	31% (2%)	231
C1D-2	Clear documentation and/or help to assist you in understanding how the data was derived and what it represents	30% (4%)	263
C1D-5	Consistency of data fields and data structure that allow linking data from multiple systems/sources	27% (5%)	190
C1D-6	Access to analytical tools to help with analysis of data (e.g., statistical software, database applications)	25% (4%)	166

**•Note:**  
 "Not sure" and "not applicable" have been excluded from analysis. % indicates respondents who evaluated the attribute excellent or good. Number in parentheses is % who scored the attribute excellent.

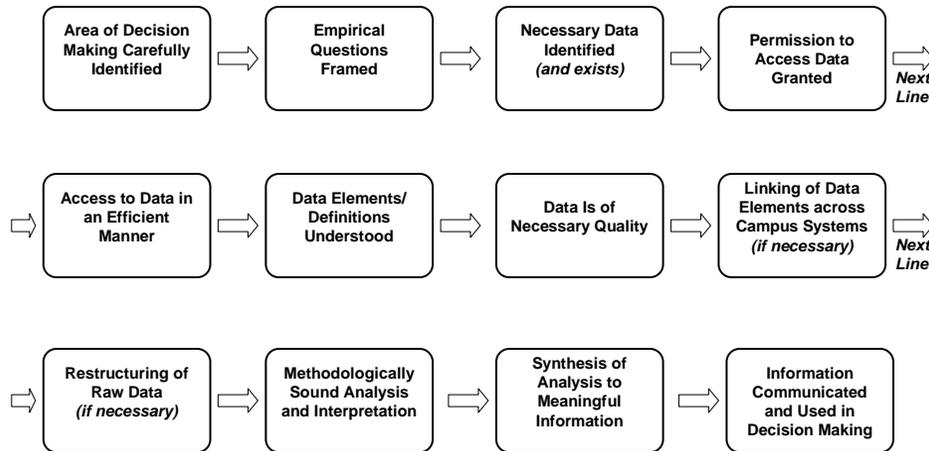
**Observation: Importance of High-Quality Analytical Work.** *Clearly, the existence of high-quality data and its security are of critical importance; but without high-quality analysis-related work, Berkeley may be severely limiting its ability to profit from that data.*

## Assessment of Campus Data Environment — Specific

*Context: Data Work Flow*

Figure 5 below helps to illustrate the potential complexity of generating useful information to support campus decision making. As illustrated in the figure, the process begins with careful specification of areas of decision making and empirical questions to be addressed; it then proceeds through a number of phases, from the collection of raw data to the development of meaningful information that can inform decision making. The bulk of these interior phases is pre-conditioned on the preceding phases. For example, the existence of necessary data is a precondition for all subsequent phases of the process. So, too, is access to the data—if data access is blocked or limited, the probability of eventually developing useful data-derived information is low at best. If data is of poor quality or questionable accuracy, or all the necessary data elements can not be linked in a valid manner, the endeavor is almost certain to fail. As the saying goes: Garbage in, garbage out.

**Figure 5: Data-Informed Decision Making Flow Chart**



The work flow from framing empirical questions of critical importance to the Berkeley campus to successful generation of useful data-derived information is frequently long and complex, and can be impeded at one or more points along the way. By determining the specific barriers to this efficient data work flow, we can begin to identify promising solutions to our current problems.

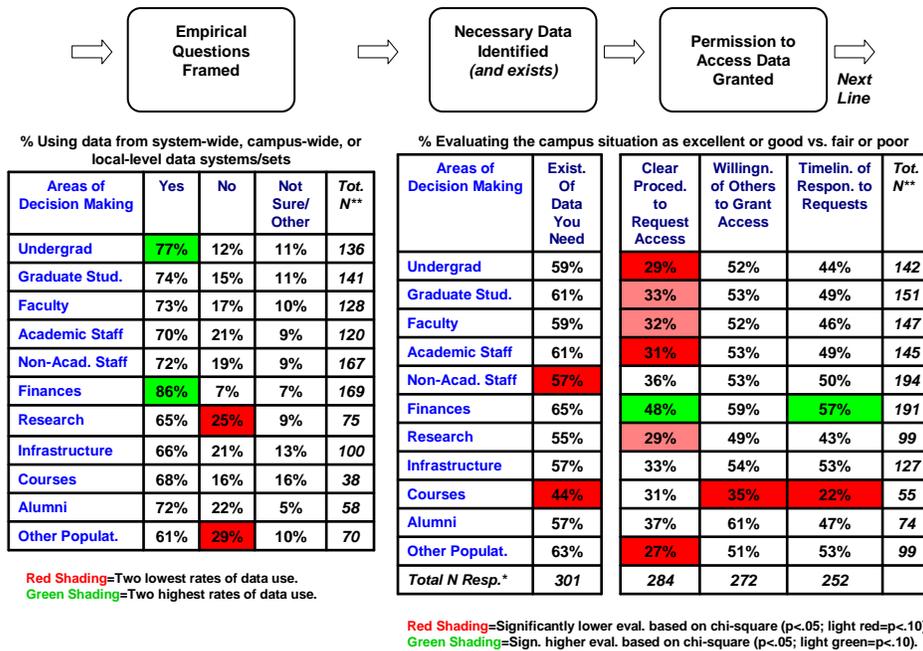
The process of deriving useful information from record-level data—frequently stored in complex transactional databases as opposed to planning/analytical databases—can be labor intensive; proposed projects should be scrutinized to assess their likely impact on the efficiency of our data-to-information flow.

***Recommendation: Prioritize Cost-Effective Projects that Increase Data Efficiency.*** Careful analysis of 1) the likely input (resources) necessary to succeed at developing useful information, and 2) the eventual yield in doing so should be undertaken early on in the process. Furthermore, the development of resources or technologies that make [the data portion of the process] more efficient for the broadest feasible array of mission-critical decision-making areas should be prioritized.

*By Major Areas of Decision Making*

The survey data enable us to begin to map the highs and lows of our campus data environment by major areas of decision making. *Figures 6a, 6b, and 6c* display an overview of respondents’ evaluation of the campus data environment by major decision areas, mapped on to the preceding work flow diagram (see *Figure 5*). *Figure 6a* shows rates of data use in the decision-making process by major areas of decision making, and the specific evaluations among campus respondents with regard to existence of data and permission to access data by major areas of decision making/support. (Note: The residual category of “other areas” is excluded from the table since it is difficult to interpret without careful review of verbatim comments). *Figures 6b* and *6c* show campus respondents’ ratings by major areas of decision making for the middle and later portions of the flow diagram.

**Figure 6a: Data-Informed Decision Making Flow Chart, Evaluation of Campus by Major Areas of Decision Making**



With regard to data use in the decision-making process (*Figure 6a*), respondents who make or support decisions in the undergraduate and financial area 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. These higher rates of data use in the fiscal (86%) and undergraduate (77%) arena are perhaps not surprising given the necessity of data systems to support the flow of money across the campus and assure fiscal responsibility, and because of the centralized nature of many aspects of undergraduate education—from application to graduation—particularly in

comparison to the decentralized nature of many aspects of graduate education and management of employees, hiring and promotion. So, too, the lower use of data with regard to the research grants/contracts arena and other populations is probably not surprising given the lack of essential day-to-day core business practices mandating immediate and constant access to centralized systems, and the fact that these other populations are not salient to many campus departments/units. Beyond these patterns, another clear implication of the data shown in this table is that **a good number of respondents are not using data from campus or local systems in support of the decision-making process, or are unsure whether data is being used in support of the endeavor.**

Based on respondents' comments, a wide range of different reasons help to account for the failure to use data fully in support of the decision-making process on the Berkeley campus:

*[T]here are not clear data sets and tools for using the data. [T]oo much localized control of what ought to be strategic university data limits broader use of data-based analysis.*

*Often takes considerable lead time to extract and then manipulate data needed as context for decision-makers. More often than not, getting the data means knowing who to ask, and answers may be different depending on the contact.*

*Data are in scattered locations. It can take a good bit of searching.*

*Not sure about this question because the many areas that I oversee have such different sources of information.*

*[T]oo much information needed from units that are less than forthcoming with the data.*

*There should be more staff available to work on data integrity/data analysis.*

***Observation: Decisions without Data Support.*** *The failure to use data or lack of clarity regarding possible use of data is of concern.*

***Recommendation: Investigate Why Data Is Not Used.*** *As much as possible, decision making should be supported with data-derived information. A concerted effort to investigate further what accounts for this (decisions made without strong data support) should be pursued, starting with a review of verbatim survey comments.*

Figure 6b: Data-Informed Decision Making Flow Chart, Evaluation of Campus by Major Areas of Decision Making (continued)



Areas of Decision Making	Release Of Up To Date Data	Availability of Data When You Need it	Ease Of Extracting/Accessing Data	Clear Documentation Re. Data	Accuracy Quality Of Data	Consistency Of Data Across Campus	Consistency Of Data Fields Across Systems	Tot. N**
Undergrad	41%	47%	40%	26%	67%	33%	23%	142
Graduate Stud.	49%	51%	43%	27%	64%	37%	24%	151
Faculty	46%	50%	44%	30%	60%	38%	21%	147
Academic Staff	48%	48%	43%	33%	63%	40%	18%	145
Non-Acad. Staff	50%	50%	44%	31%	69%	36%	27%	194
Financial	54%	53%	49%	38%	69%	43%	29%	191
Research	38%	44%	40%	33%	58%	34%	27%	99
Infrastructure	48%	45%	39%	33%	64%	32%	25%	127
Courses	28%	48%	21%	23%	56%	33%	14%	55
Alumni	49%	50%	30%	31%	62%	33%	25%	74
Other Populat.	48%	45%	41%	41%	66%	41%	30%	99
Total N Resp.*	266	240	265	263	263	212	190	

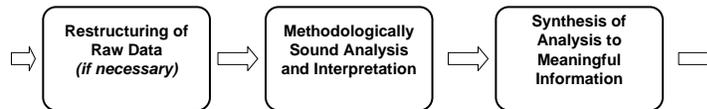
Red Shading=Significantly lower evaluation based on chi-square (p<.05; light red=p<.10).  
 Green Shading=Sign. higher evaluation based on chi-square (p<.05; light green=p<.10).

The remaining tables included in Figures 6a, 6b, 6c display the percentage of respondents who evaluated a particular aspect of the campus data environment as excellent or good, vs. fair or poor, by major areas of decision making. Specifically, the ratings of respondents who indicated that they were involved in a particular major area of decision making on the campus are displayed by rows. Thus, an individual's rating of a particular aspect of the campus data environment can show up in multiple rows, assuming she or he participates in multiple major areas of decision making. The dark green shading denotes that respondents who work in a particular area rated a particular item significantly higher than respondents who do not work in this area (significant at the .05 level). We also include light green shading that denotes significant differences at the .10 level, which is frequently considered to be marginally significant at best. The red shading indicates that respondents participating in a particular area rated the item lower (dark red shading, p<.05; light red shading, p<.10).

Since individuals were asked only once to globally assess the campus environment with regard to each particular issue, and yet many individuals work in multiple major and sub-topical areas, the global ratings that are displayed by each major area are likely to reflect respondents' evaluation of both this area and any other areas of decision making in which they work. Accordingly, significant differences by broad decision-making areas (and later sub-topical areas that will be discussed in the latter portion of this section) are in general likely to understate the degree of the difference by major areas and sub-topical areas. This aspect of the study design along with the low counts for some sub-topical areas of

inquiry explains why we have shaded items at the .10 level in addition to the more traditional .05 threshold.

Figure 6c: Data-Informed Decision Making Flow Chart, Evaluation of Campus by Major Areas of Decision Making (continued)



Areas of Decision Making	Access To Standardized Reports	Access To Specialized Reports	Access To Trained Staff	Access To User Friendly Reporting Tools	Access To Analytical Tools to Help with Data	Tot. N**
Undergrad	34%	23%	40%	25%	29%	142
Graduate Stud.	35%	28%	48%	28%	23%	151
Faculty	37%	29%	46%	27%	25%	147
Academic Staff	39%	32%	47%	31%	28%	145
Non-Acad. Staff	42%	32%	41%	33%	22%	194
Financial	52%	38%	46%	36%	19%	191
Research	50%	35%	43%	22%	18%	99
Infrastructure	46%	31%	41%	32%	22%	127
Courses	24%	19%	41%	19%	13%	55
Alumni	35%	30%	44%	26%	15%	74
Other Populat.	42%	30%	41%	26%	17%	99
Total N Respondents*	248	217	244	231	166	

Red Shading=Significantly lower evaluation based on chi-square ( $p < .05$ ; light red= $p < .10$ ).  
 Green Shading=Sign. higher evaluation based on chi-square ( $p < .05$ ; light green= $p < .10$ ).

Despite the constraints of the displayed data, certain patterns observed in these figures appear to be of compelling interest—particularly as they are compared to verbatim comments and in-depth case studies that constitute a portion of the Advisory Group’s larger effort. Most notably, respondents who make or support decisions regarding financial information on the campus are significantly more likely to evaluate positively three particular items—clear procedures for requesting access to data, timeliness of response to data requests, and clear documentation—than respondents who do not work in this area. And in the case of all of the items presented in *Figures 6a, 6b, and 6c*, **individuals working on financial issues display either the highest or second highest overall rating for each displayed item.** Clearly, in the aggregate, this sector of Berkeley decision making is the most favorably evaluated. In contrast, individuals who work on course-related areas rate five particular items at a significantly lower level: existence of data, willingness of other units to grant access to data, timeliness of response to requests for data, release of up-to-date data, and ease of extraction of data. Thus, **individuals working on course-related data seem particularly distressed by the current Berkeley data environment.**

**Observation: Highs and Lows of Campus Ratings.** 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.

A number of respondents discussed their difficulties dealing with course-related data on the campus:

*The data model for courses is nearly unusable due to the simple fact that courses are NOT referred to by unique identifiers. Further, the numbering system we DO use for ID has been misused in repeated efforts to model new attributes in existing elements.*

*Hindered by the campus' inability to track courses taught by faculty/dept. over time. We're forced to maintain an in-house, unwieldy spreadsheet when campus in fact has all the data via the course scheduling system.*

*Frequently hear complaints that databases in course scheduling and in student services are difficult to extract useful data from, and frequently requests have to be made of personnel directly supporting those systems to retrieve needed data.*

*The user interface for reviewing and updating [course] data is archaic. Many staff dread the task of reviewing and correcting this data and it is common knowledge that the staff who are responsible for the actual data entry oftentimes ignore or manipulate the ways of reporting so that the problems associated with correcting errors and omissions report data will go away.*

*[The dataset Course] is not in a form that it can be readily accessed by anyone outside the Academic Senate. Still on mainframe.*

*I really don't understand the amount of money and time that has gone into DARs [the degree auditing reporting system] and yet it still does not work. Seems to be good money after bad.*

**Observation: Necessity of Detailed Understanding.** Preserving and enhancing excellence in the University's myriad data efforts requires a detailed understanding of the campus's overall data landscape.

**Recommendation: Continue Mapping and Analyzing Campus Data Landscape.** As the University's organizational structures, priorities, and data needs evolve, ongoing mapping and

*analysis of the campus data landscape will be necessary if we are to preserve and enhance excellence in our overall data efforts.*

A summary review of *Figures 6a, 6b, and 6c* also suggests that one particular area of data flow is causing an undue bottleneck in a number of different decision-making areas, clear procedures for requesting access to data.

***Observation: Importance of Clear Procedures to Request Data Access.*** *Clear procedures for requesting data access appears 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, research, and other populations.*

As stated above, the overall campus rating for this item is quite low, with fewer than two in five respondents evaluating the campus favorably in this respect; the financial area is the only bright light with regard to major areas of decision making. Release of up-to-date data is also a particular problem in regard to undergraduate issues, research, and courses.

***Observation: Importance of Access to Data.*** *Access to high quality data is essential to supporting informed decision making on the campus.*

***Recommendation: Port Successful Procedures to Problem Areas (if appropriate).*** *Since the financial sector of campus decision making has demonstrated the viability of having relatively successful procedures regarding data access issues, the possibility of porting these methods over to other areas of decision making should be explored.*

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

Although the broad findings discussed above are useful, the nuances of the Berkeley data landscape can be better mapped by examining overall ratings by sub-topical decision areas, access to campus datasets (actual or requested), and job type. *Figures A-3—A-20* in the *Appendix* show significant differences ( $p < .10$ ), positive and negative, for the assessment items of greatest concern to this particular effort (some of the security-related items are not shown in the *Appendix*). This mapping provides a clear view of our current peaks and valleys. A brief discussion of salient patterns follows.

**Existence of Data You Need.** Nine out of the 13 sub-topical areas of decision making that were associated with significantly lower overall campus ratings were in the broader category of non-academic staff (see *Appendix, Figure A-3*), a pattern previously observed in *Figure 6a*. Clearly, a good amount of this staff data already exists in HRMS, but many respondents seem to not be aware of this; and accessing data to support decision making at the necessary level of aggregation may be difficult, given the relatively limited reporting functions currently available for HR data in BAIRS.

The fact that policy analysts, systems programmers, and Institutional Research (IR) analysts are more likely to rate this item relatively favorably is probably not surprising, given the fact that knowing about the existence of data and making use of it is frequently an important functional responsibility of positions of this type. (It has been often observed in IDMG-related meetings that, for the most part, data exist, but it is the subsequent parts of data informed decision-making work flow—see *Figure 5*— where most difficulties arise). This pattern is also likely reflected in the higher ratings on this item by individuals who have access to Office of Student Research Survey Data Bases and the Student Data Warehouse, since the above job groups are particularly likely to have access to one or both of these systems.

***Observation: Existence of Data Is Not Our Primary Problem.** Because the job groups who are arguably among the best positioned to assess the existence of necessary data on the campus rate this item more favorably than most other job groups, it is reasonable to conclude that 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-Informed Decision Making Flow (22) 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*

***Recommendation: Focus Improvements on Increasing Data Access and Consistency.** Because the middle portion of the Data-Informed Decision-Making Flow (22) is overall most encumbered, future efforts should prioritize access to and consistency of data to allow for increased production of high-quality analysis.*

**Established Procedures for Requesting Access to Data.** Centralized campus systems, particularly ones related to financial data, are associated with more-favorable ratings on this item than

are course-related systems and some undergraduate systems and sub-topical areas (see *Appendix, Figure A-4*). Presumably, as mentioned previously, effective procedures have been established for accessing these particular large-scale centralized systems (if so, the relative success of these approaches offers the possibility of utilizing the methods in other areas). Although campus-level decision makers and staff who support them rate this item higher, staff members who support departmental chairs give it at a significantly lower rating.

**Willingness and Timeliness of Other Units to Grant Data.** Individuals who work on course-related decision making (e.g., approval/modification of courses, course offerings/profiles) rate both of these items more negatively than others (see *Appendix, Figure A-5—A-6*). The admissions database, BIBS, CADS, and Cal Profiles are associated with higher positive ratings in regard to willingness to grant data while the UGA Admissions system and FASDI received lower ratings; but individuals who support chairs and systems managers rate this item more negatively. And space and financial-related areas of decision making and data systems are associated with more positive ratings in regard to timeliness of cross-unit response to data access.

***Observation: Inadequate Access to Data for Those Who Support Academic Chairs.** Gaining access to necessary data appears to be particularly problematic for staff who support departmental chairs. If the staff of departmental chairs are blocked from access to necessary data, departmental-level decision-making activities are likely to be compromised.*

***Recommendation: Address Accessibility Issues on the Departmental Level.** Further investigation should be undertaken to alleviate any potential bottleneck in this regard, particularly in light of the fact that departmental chairs are involved in such a large number of campus decision-making areas but appear in general to rate the campus situation less favorably than most other job groups.*

**Release of Up-to-Date Data.** A large number of sub-topical decision areas are associated with lower ratings, many of which fall into the broader areas of course-related issues, research, faculty, academic staff, non-academic staff, and undergraduate issues (see *Appendix, Figure A-7*). Many of these topical areas are complex and require access to large amounts of data (e.g., mentoring and climate issues, graduation and productivity issues, longer-term planning issues, etc.). Clearly, individuals who

have access (or have requested access) to the campuswide financial systems rate this issue more favorably, as do general office analysts. Deans, in contrast, rate this item less favorably.

***Observation: Core Business Functions vs. Planning, Analytical, and Assessment Functions.***

*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. Certainly, core daily business needs (transferring funds, budget accounting, hiring employees) are of critical importance to the campus; so too, however, are broader-scale planning issues, and human-centric areas that directly relate to recruitment, retention, and productivity of employees and students. In general, the campus appears currently stronger with regard to meeting the immediate needs of core business functionality, and weaker with regard to planning, analytical, and assessment functions, including those involving human-centric climate issues.*

***Recommendation: Improve Deans' Access to High-Level Data.*** *Since deans are frequently charged with dealing with non-business planning, analytical, and assessment issues, their tendency to rate the campus poorly with regard to release of up-to-date data should be addressed.*

**Availability of Data When You Need It.** Many sub-topical areas related to staff are associated with lower ratings on this item (see *Appendix, Figure A-8*). Some respondents expressed frustration with limits to the hours that various campus systems are currently available, particularly systems related to employee data: “HRMS shutting down at 6pm is too early”; “The main data systems that I use BFS/BAIRS and HRMS have a 6pm cut-off on weekdays and is totally unavailable on some parts of the weekend”; “It would help if HRMS/BAIRS were available after hours and weekends.” In contrast, access to OSR survey data, the admissions data base, graduate student information systems, and Cal Profiles is associated with higher ratings. All of these campuswide systems provide some level of Web access that allows users to access data directly whenever they need to from any computer with Internet access, on or off campus).

**Ease of Extracting/Accessing Data.** A large number of sub-topical areas of decision making are associated with lower ratings on this item, particularly with regard to alumni or development issues,

non-faculty academics, and non-academic staff (see *Appendix, Figure A-9*). BIBS and BIS are associated with higher ratings, as is the topical area of budget development and allocation. Respondents who need to access Student Information Systems and BearFacts rate this item less favorably.

Given these patterns, the ongoing requests to update alumni/development systems and students systems appear to be in part a response to the current perception that accessing data is a cumbersome and inefficient process. Since staff data are already centralized in HRMS, the limited attention given to reporting and data access interfaces may account for their association with more negative ratings.

**Clear Documentation Regarding Data.** Financial budget sub-topical areas and some of the campuswide financial systems are associated with more-favorable ratings on this item (see *Appendix, Figure A-10*). In contrast, student information systems, the Student Data Warehouse, graduate information systems, CSIR, CARS, and BearFacts are associated with very low ratings (under 15% of respondents rated the campus situation as excellent or good).

**Accuracy/Quality of the Data.** A number of sub-topical areas are associated with lower ratings on this item, particularly with regard to faculty and research activities (see *Appendix, Figure A-11*). Certain financial systems are, however, associated with higher ratings, including the Cashiers Deposit System, BIBS, and Payroll. As some respondents pointed out, data accuracy is dependent on the quality of what is put into the data systems: “Hard to judge the quality, since it is only as good as what is entered by people across the campus”; “[B]ecause we are decentralized and dependent on many others to enter data, the data we utilize in eRecruit particularly is inaccurate and incomplete”; “[C]ampus departments are not inputting data (incomplete) and sometimes not accurately.”

***Observation: Possible Danger to Sound Decision Making.** Although the campus in general received relatively more-favorable marks regarding accuracy and quality of data than many other items, one or more sub-topical areas are associated with a lower rating (e.g., below 50% excellent/good). These present a danger of or perception of poor-quality or inaccurate data compromising the decision-making process.*

***Recommendation: Investigate Perceptions of Data Inaccuracy and Mitigate Any Identified Problem.** Any sub-topical area that is associated with a lower rating (e.g., below 50% excellent/good) should receive further investigation to determine whether there are inaccuracies in the data with an eye to improvements. If data are accurate but a perception of inaccuracy exists, investigate what accounts for this perception and seek to mitigate it.*

**Consistency of Data across the Campus.** Sub-topical areas associated with alumni/or development issues are associated with lower ratings on this item (see *Appendix, Figure A-12*). BIBS and Cal Profiles are associated with higher ratings, and data recorders rate the item the most favorable of all job types. The fact that the Student Data Warehouse is associated with lower ratings might, in part, reflect the challenges that were encountered by participants in this project with regard to campuswide data consistency issues, or the lack thereof.

**Consistency of Data Fields across Systems.** All of the significant differences on this item fall on the negative side; and, as previously noted, the overall campus rating is particularly bad (see *Appendix, Figure A-13*). Systems managers and policy analysts rate this item even lower than the other job types do.

***Observation: Poor Consistency of Data Fields across Systems.** The fact that systems managers and policy analysts—who likely possess significant expertise in this area—are particularly likely to rate consistency of data fields across systems in the negative suggests that this is a particular area of concern for the campus.*

***Recommendation: Make Consistency of Data Fields across Systems a Campus Priority.** The lack of consistent data fields across campus systems and clear definitions undermines our ability to conduct high-level analysis, support well-informed decisions, and represent ourselves in a consistent and clear manner. The campus as a whole needs to prioritize consistency of data fields across existing systems and in future efforts to improve the data landscape.*

**Access to Standardized Reports.** Certain subtopical areas under the broader categories of research/contracts and finance are associated with higher ratings, whereas sub-topical areas related to undergraduate and graduate students are associated with lower ratings (see *Appendix, Figure A-14*). Course-related systems are associated with lower ratings, as compared to a number of the financial systems that are associated with higher ratings.

**Access to Specialized Reports.** The patterns are largely similar to those above, except that graduate student sub-topical areas are not associated with negative ratings, and the financial systems, except for BFS, are not associated with more-positive ratings (see *Appendix, Figure A-15*).

**Access to Trained Staff.** Given the small number and nature of the significant findings with regard to this item, there is little that can be gleaned from them, except perhaps that members of the Chancellor's Cabinet seem to have better access to trained staff (see *Appendix, Figure A-16*).

**Access to User-Friendly Reporting Tools.** As noted previously, this item 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 (see *Appendix, Figure A-17*). The fact that institutional research analysts, policy analysts, and staff who support campus-level decision makers or deans rate the campus more favorably in this regard (albeit still largely toward the negative side) suggests that there may be more user-friendly reporting tools of which others may not be aware (or the former group's criteria for what is user friendly may be more lenient).

**Access to Analytical Tools to Help with Data.** A spattering of subtopical areas are associated with particularly unfavorable ratings on this item, but a couple of high points jump out: OSR's survey systems and research databases (see *Appendix, Figure A-18*). OSR has devoted considerable energy and demonstrated considerable adeptness at building tools to support data analysis (certainly these should be considered as a possible model for other areas where the data landscape is more negatively evaluated). A similar pattern, as noted earlier, is seen with regard to job type, with IR analysts, policy analysts, and, in this case, systems programmers evaluating the campus more favorably on this item, and campus-level decision makers rating the item particularly low (only one out of 16 individuals rated the situation as excellent or good). This pattern may reflect either a greater knowledge on the part of IR/policy analysts and systems programmers with regard to available analytical tools, a greater capacity to use available analytical tools, or, most likely, a combination of both.

***Observation: Unequal Knowledge and Access.** 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. Indeed, some respondents not in these analyst groups noted in the survey that they had no idea that so many data systems existed on the campus.*

***Recommendation: Disseminate Inventories of Data and Analytical Resources.** The campus should develop and effectively disseminate clear and easily digestible inventories of existing data and reporting tools.*

**Procedures to Protect Data.** While sub-topical areas under the broader areas of alumni/development and graduate student issues rate more favorably, sub-topical areas under non-academic staff rate less favorably and seem to call for some level of additional investigation. SAMS and Student Information Systems are associated with particularly high ratings, making them possible models for other areas (see *Appendix, Figure A-19*).

**Procedures to Ensure the Proper Sharing of Data.** In contrast to many of the preceding areas of the data landscape, all of the sub-topical areas with significant differences from the other areas fall on the positive side of ratings, with alumni/development and space sub-topical areas associated with particularly positive ratings (see *Appendix, Figure A-20*). As a group, deans rated this item on the favorable side, whereas campus-level decision makers and members of the Chancellor's Cabinet rated it on the less favorable side.

***Observation: Access to Data vs. Security of Data.** Though there is an inevitable tension between strong data security and ready access to data, both are essential to furthering Berkeley's mission. At present, the campus is rated more favorably for securing data than providing access to data.*

***Recommendation: Improve Access to Data While Maintaining Security.** As we move forward, efforts to increase access to data should be prioritized while security of data maintained.*