Our current approach to the governance of data as a campus resource has created a situation that is unsustainable. Data is replicated over and over again, modified and fragmented by the many independent systems and the organizations who manage them. Security weaknesses exist as a result of both the spread of the data and the vast number of varying architectures used to collect, distribute, and manage the data. The administrative and academic units that manage duplicated data (or create duplication for the purpose of local management) all assert some level of ownership to each data element. Considerable and overlapping investments continue to be made across campus silos with technologies and staff time for both functional and technical groups to maintain this “spaghetti architecture” of legacy, current, and emerging technologies.

This represents a divergent strategy at a time where, more than ever, the campus needs greater visibility, accessibility to, and sharing of data within and across silos to make better decisions and work together as a campus. The opportunity here is not only for administrative data, but also for scholarly information. With the advent of new technologies and as new web services become available we will see new ways in which scholarship can benefit from breakdowns of information silo’s, enabling interdisciplinary research, teaching and greater access to the public. Examples include access to theses, dissertations, lab data and faculty research, collections from our museums and many others.

A pressing example of the immediate governance challenge we face is development of the Campus (Enterprise) Data Warehouse (EDW). Starting from a foundation of financial and human resources data developed for the BAIRS reporting portal, we have studied the opportunities of having a truly campus-wide data warehouse representing subject areas beyond finance and HR, and proposed an architectural solution in reports published to the campus. However from the inception of the studies nearly 2 years ago, it was universally recognized that the core decision making structures necessary to support prioritization and commitment beyond existing data was absent at a campus level. We have since made little progress in developing a comprehensive implementation plan for the EDW, let alone moving ahead with implementation. This has contributed to parallel implementation efforts and investments with separate data warehouses and reporting systems for use across the same university community.

The issues involved are complex, however, inadequate decision making and the pace of that decision making equates to missed opportunities and greater risk.

Requests for data across organizational boundaries are increasing dramatically. However, we have not adequately invested in the tools, infrastructure, processes or organizations and staffing needed to make the data safely available in an environment which requires careful management of the conflicting goals of openness and protection.

The data itself may be sensitive with “ownership” or stewardship issues associated with it. While we have no perceived shortage of policies and regulations to direct how this data should be managed or consumed, we do lack a compliance program to measure how effectively the policies and regulations are being implemented. Compliance as proactive risk management is
performed as a “best efforts” approach rather than a systemic program, the implementation of which would - perhaps - make data proprietors feel easier about their data being shared.

To address the above set of needs at an institutional level, we believe the implementation of a new governance framework is necessary to optimally and strategically manage the investments made on data-centric activities and manage risk. Broad campus representation is needed to a) speak to and align university’s priorities; b) advocate funding for initiatives that support these priorities; c) ensure that the initiatives and related services are being delivered appropriately to the campus stakeholders.

Data Management Governance Framework
The new governance framework will include the following:
Refer to Figure 1 – Framework for Campus Data Governance on the following page.

1. Data Management Steering Committee
   - Four to six executive data proprietors who meet biannually to provide the overall objectives and to support funding strategies for the various data initiatives based on campus strategic goals

2. Data Management Leadership Council
   - Ultimately, 10-14 senior officers of the campus meeting at least quarterly to represent various departments across the campus to oversee a broad portfolio of data initiatives and services. This group will likely begin with five to seven individuals with a cross section of data domains represented. Beyond administrative interests, this council will require significant representation from the research and faculty (Academic Senate) communities as well as the Libraries/Museums. This is not an advisory group, but a team assembled with formal responsibilities to assign priorities and funding, approve initiatives and monitor the performance of sub-programs and the delivery of related projects and services to the campus. This group would also be tasked with mediation of data stewardship issues as they arise.

3. Program Managers
   - We have identified a number of programs of critical importance which will evolve and expand based upon the direction of the Leadership Council.
     I. Campus (Enterprise) Data Warehouse and Business Intelligence (EDW/BI), to address the development of a central data reporting and analysis environment which allows examination, planning and analysis across campus data silos.
     II. Data Integration, to address transaction or event based integration of structured and unstructured data between campus applications. This is inclusive of traditional ETL (Extract, Transform & Load – batch oriented) and EAI (Enterprise Application Integration – transaction-based) methods. Initially this role will be included in scope of the EDW/BI Program Manager.
     III. Metadata Management – to address the management and publication of information about the data necessary to make it understandable and contextual. Initially this role will be included in scope of the EDW/BI Program Manager.
     IV. Repository and Content Management – to address management, accessibility and integration of unstructured and structured data e.g. documents and digital media managed in media vaults and other repositories.
V. **Policy and Compliance** – formally managing the standards and policies related to the management and use of data across the campus, but with the addition of a compliance function which measures the effectiveness of the policies to implementation, with associated risk management

- Program managers should be established over the coming years to represent various overlapping dimensions or sub-programs of campus data management. They will be primarily responsible to manage the incremental growth of capabilities of their programs, consistent with the budgets and priorities established by the Leadership Council. The Program managers will manage “Centers of Excellence” that will facilitate and monitor the performance of various campus delivery teams (below) who are responsible for program strategy, the portfolio of development projects and service operations. The “Centers of Excellence” are expected to foster best practices, facilitate education and training of their area.

4. **Delivery Teams for each program**

- Virtual teams (from multiple campus units) made up of functional campus representatives and technology service providers who perform the following functions:
  
  i. **Strategy** – defining future state data, application, technology and process architectures to support campus priorities. Program managers and strategy teams for all programs will be in regular discussion due to their inherent interdependencies.
  
  ii. **Development** – project-based development of capabilities as defined by the strategy team using consistent methodologies, technologies and practices
  
  iii. **Operations** – deployment and support of new increments of capabilities to the campus community
Specific proposal for FY07-08:

The new Governance structure proposed here is ambitious and complex (as is needed by the Berkeley campus for its data management). It will likely take 3-5 years to build out a full structure that supports the both administrative and academic opportunities and challenges. We are recommending the following steps be taken in FY 07-08 with the corresponding budget request. The new Governance team established here will then further define the road forward and plans and budget required for future fiscal years. Note, too, that as part of 07-08 actions the new governance team would reformulate the role of the existing Data Stewardship Council.

Steps and funding for 07-08:

a) The formation of Data Management Steering committee, who will be tasked to appoint the Data Management Leadership Council
b) The appointment of the Data Management Leadership Council to assess campus data priorities, and the appointment of a half-time (50%) resource to act as staff for the council.
c) The appointment of a full time Program Manager for the Campus (Enterprise) Data Warehouse / Business Intelligence. In the short term, this program manager would also oversee integration and metadata management programs until expansion is warranted.
d) The appointment of a Program Manager staffed half-time (50%) to formalize a Repository and Content Management Program, and will focus in FY07-08 on supporting media vault and other repository services to the campus academic community.
The appointment of a Program Manager staffed half-time (50%) to formalize a Data Policy and Compliance Program.

1. **Alignment with IT Strategic Plan**
   This proposal directly relates to Critical issue 5: Governance, funding, structure. However, because this proposal addresses investment in data resources across administrative, academic (research and teaching) and public boundaries, the proposal also aligns with Critical issue 1: Teaching, learning; Critical issue 2: Student experience, from prospects through alumni; Critical issue 3: Research; Critical issue 4: Security, reliability, access.

2. **Impact**
   The entire campus and UC system has a vested interest in ensuring data resources are managed securely, in accordance with appropriate policies and regulations, and at the same time making the data accessible to those administrators, academics and students who will benefit from strategic decision making, and operational agility and effectiveness.

3. **Risk assessment**
   If we do not adopt a new data governance approach, decisions and commitments made around data will continue to be made at a tactical level, in silos without true campus alignment. Decisions on strategic investments will be too slow for the campus to realize major opportunities, manage risk, and to ensure investments are made in an optimal way. The EDW continues to be an example of this.

   Acceptance of this proposal should be performed with consideration for potential overlaps with existing campus groups such as the Campus Technology Council, Data Stewardship Council and the Academic Senate. We see the activities of this governance framework to be complimentary to the efforts of these groups and therefore boundaries and supporting processes need to be defined. Opportunities may exist to share staff among the groups.

4. **Innovation**
   We are convinced that the complex data issues experienced by UCB are shared with growing realization by many other similar universities nationally and internationally. The University has recently shown great innovation with the creation of a campus technology governance structure, and with the organization design of IS&T, and specifically the formation of a Data Services group. UC Berkeley could be one of the first universities nationally to address the emerging complexities of managing administrative and academic data with a robust governance structure that is integrated across the institution in this way.

5. **Funding model**
   We seek funding for the following positions for FY07-08:
   
   - 0.5 FTE - DM Leadership Council Staff (1 FTE FY08-09)
   - 0.5 FTE – Program Manager of Repositories and Content Management (1 FTE FY08-09)
   - 0.5 FTE – Program Manager of Data Policy and Compliance (1 FTE FY08-09)
   - 1 FTE – Program Manager of EDW and BI

   Activities delivered through the programs themselves should be appropriately defined in terms of scope and cost to the Program Manager, who should then approach the Leadership Council with a separate funding proposal. The Leadership Council should review the merits of the request, and seek sponsorship from the Steering Committee and secure funding through the Steering Committee and CTC.

   In future years, we see the DM Leadership Council as the body who would propose budget requests through the CTC process for data related initiatives.
### Summary of Funding Model: CIO: Data Management Governance

<table>
<thead>
<tr>
<th></th>
<th>FY 05-06</th>
<th>FY 06-07</th>
<th>FY 07-08</th>
<th>FY 08-09</th>
<th>FY 09-10</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Campus funding sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporary/development</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
<tr>
<td>Permanent/on-going</td>
<td>$0</td>
<td>$0</td>
<td>-$311,775</td>
<td>-$496,057</td>
<td>-$510,639</td>
</tr>
<tr>
<td><strong>Total campus funding sources</strong></td>
<td>$0</td>
<td>$0</td>
<td>-$311,775</td>
<td>-$496,057</td>
<td>-$510,639</td>
</tr>
<tr>
<td><strong>Other funding sources</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Temporary/development</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
<tr>
<td>Permanent/on-going</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
<tr>
<td><strong>Total other funding sources</strong></td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
<td>$0</td>
</tr>
<tr>
<td><strong>TOTAL ALL FUNDING SOURCES</strong></td>
<td>$0</td>
<td>$0</td>
<td>-$311,775</td>
<td>-$496,057</td>
<td>-$510,639</td>
</tr>
</tbody>
</table>

**Note:** In accordance with the University’s accounting system, positive numbers are expenses or deficits, while negative numbers are funding or surpluses.