I. Executive Summary

In 2011, a Research Informatics Strategic Planning (RISP) process was initiated by the Office of Information Technology (OIT) and the Office of the Vice Chancellor for Research (OVCR), in recognition of the need to specifically identify current and future research directions and to anticipate the types of data that researchers will be using and how it is collected, processed, accessed, analyzed, leveraged and shared. For the purpose of this strategic plan the term research informatics is used to mean all data collection, generation, computing, storage, analyses, visualization, access and sharing in support of research, as well as supporting infrastructure, technology, software, tools, services, policies and governance.

The Need for a Strategic Plan

The need for an institutional strategy has become increasingly important. The enormous acceleration in the amount of data produced over the last century has created a shift in how research universities use data, in terms of quantity and methodology. The specific act of extracting knowledge from digital data has now become an academic subject unto itself, and researchers from across entire universities are becoming more engaged in conducting research in the emerging field of informatics. In light of these trends, it has become evident that a leading research university in the 21st century will be defined by: (1) how successful it is in recruiting faculty who are adept at generating and analyzing data; (2) the extent to which researchers across campus form interdisciplinary teams to solve grand challenges; and (3) how well the university educates future workforces on effectively using 21st century cyber infrastructure.

The timing and urgency of the RISP effort is further motivated by the following needs and opportunities: (1) UCLA has been the recipient of a number of large, research informatics oriented awards in the past 3 years; (2) the U.S. has focused national attention and investment on healthcare, manufacturing and energy, which involve complex informatics infrastructure; (3) cutting edge computing (exaflop and exabyte) is around the corner; (4) all federal agencies have organized programs around informatics and big data; and (5) there is a national emphasis, led by the National Institutes of Health and the National Science Foundation, to ensure that research data is accessible and shared with the broader research community.

Strategic Planning Process and Goals

The goals of the RISP process are as follows:

- Develop a 10-year vision for UCLA and the strategic principles that will guide planning
- Develop a 5-year detailed roadmap that defines priorities and critical paths
- Scope effort and develop a cost estimate for a 5-year initiative
- Develop a 1-year project and implementation plan
- Define the governance process for research informatics on campus
To ensure the development of a tangible action-oriented plan, the planning process was divided into three phases:

- **Phase One:** Defining research informatics needs, barriers and obstacles that currently exist for faculty. This phase was structured around 5 discipline-oriented committees involving over 70 faculty and staff. The committees identified a series of foundation level informatics services and support needs for researchers that are now being address through various means, including a Research Resource Portal.

- **Phase Two:** Identifying a vision of themes that not only set the context for the first acts but also if pursued would enable UCLA to differentiate itself as a leader in research informatics in the next 5-10 years. One-on-one interviews and small group meetings were conducted with the Institute for Digital Research and Education (IDRE), the IDRE Board, the Humanities, Arts, Social and Information Science Group (HASIS), the Clinical Translation Science Institute (CTSI) Board, and additional faculty in social sciences. An external advisory committee was also engaged.

- **Phase Three:** Creating a new governance system to ensure cross-campus coordination of the research informatics strategic plan on campus.

**UCLA’s Strategic Themes**

Through faculty discussions the following 10 categorical themes emerged:

1. **Cross-disciplinary Collaboration.** Increasing opportunities to share ideas and think about complex research and global problems together.

2. **Research Ecosystem (rEcosystem).** Establishing shared resources that extend informatics and research capacity.

3. **Enabling a broader base of researchers.** Introducing and providing access to easier-to-use tools, services, and support to researchers who may be in silos, do not have adequate resources, or are in under-developed fields.

4. **Data Ownership and Big Data.** The ability to more readily access, collect and store massive amounts of disparate data (big data) that investigators may not wholly generate or own in their own research in order to create accelerated and novel opportunities for non-traditional research analysis and decision making.

5. **Real-time Dynamic Data.** Real-time, dynamic data and analysis is transforming traditional research approaches and methodologies by accelerating the generate-analyze-apply-learn research cycle.

6. **Multi-Use Data.** An increase in multi-use data is blurring the boundaries of research, business, and operations, allowing research to more seamlessly integrate into business workflow and operations. Policy and governance is necessary to establish compliance standards for managing such multi-use data.

7. **Image Data.** Image data research will continue to grow in sophistication as the analysis of image features increases in granularity and descriptive detail, implicating changes related to institutional infrastructure.

8. **Data Visualization.** There is a need to develop tools that will organize various data sets into a coherent visual display that can be readily interpreted and understood, such as statistical visualizations, geographical information systems and 3D modeling.
9. **Citizen Scholar.** It is becoming more common for individuals that are not employed by a college, university or other research-based organization, to engage in research purely as an enthusiast, as a "citizen scholar," which in time will drive demand for access to data sets across academic disciplines. Access to data and the products of data analysis will allow broader interpretation of results and outcomes of research projects, and potentially inspire new research areas.

10. **Mobile and Social Networking.** The wireless revolution is producing an ever-widening and thickening blanket of human-centered and sensor-based data. Mobile and social media has transformative powers, and will allow faculty to form powerful connections and reach new people that previously could not be accessed from the ivory tower.

These 10 themes viewed together, and coupled with associated action and investment, make up the focal point of and drive UCLA’s informatics strategic planning initiative.

**Transformative and Catalytic Recommendations**

The following transformative and catalytic recommended actions also emerged from the planning process as a way to support the 10 cross-campus themes. These actions are designed to set into motion a series of directional changes for UCLA that address an institutional strategy and build institutional capacity where most useful, while also recognizing the distributed and domain-driven nature of UCLA research.

**Transformative Recommendations**

- Establish and invest in a campus-wide Informatics Scholar Program in which advanced undergraduate, graduate, and post-doctoral students work with domain scholars and/or computer science, statistics, and applied math faculty on research projects across research informatics that are strategically selected.
- Establish workshops for non-health sciences faculty and graduate student researchers to learn and discuss UCLA health data marts available for analytical use.
- In partnership with IDRE and CTSI, create a shared bioinformatics resource pilot focused on genomics to investigate the potential for shared experiences, expertise and tools as well as to define the network, computational, storage and software infrastructure to support it.
- Provide a base level of no cost, managed data storage (in this context, data also includes software) in combination with enacting research data sharing plans to meet a pervasive need and to incentivize faculty researchers to: (1) register their database content; (2) provide metadata and documentation associated with their data set so that it is discoverable and usable; and (3) store backup data in one of several campus managed data storage services that meet standards for their discipline/department and for sharing/accessibility. This will also permit monitoring the frequency of data downloads and access and provide a mechanism for giving credit to faculty for generating and/or managing highly visible data bases.

It is expected that the above recommended actions will: (1) dramatically reduce barriers for faculty in using computational science; (2) facilitate access to existing data sources; (3) allow
UCLA to stay in the forefront of "big data" analytics; (4) foster collaboration and significantly broaden faculty participation in data driven research; (5) integrate education in research informatics in cohesive way on campus; and (6) bring visibility to UCLA in “big data.”

Catalytic Recommendations

- Institute a FacTech program for research informatics and take advantage of already existing FacTech department programs which function as faculty/department-sponsored ‘speed dating’ events that are customized for discipline-specific interests and quickly introduce faculty to relevant tools and practices.
- Establish and invest in an Informatics Sabbatical Program that recruits outstanding informatics scholars to spend sabbaticals at UCLA and which allows UCLA faculty to spend sabbaticals working at leading institutions in informatics. This program will also function as a ‘One-Stop Shop’ for researcher consultation on campus resources, policies, processes, training and capacity-building in combination with a web-based information and resource clearinghouse.
- Instantiate a Symposium Series for UCLA museums and campus departments on image archiving and the standardization and use of image metadata.
- Run mobile technology and social networking research innovation application contests where contestants develop mobile and social networking research applications that can be repurposed and used by other researchers for similar applications.
- Establish and invest in big data analytic workshops and summer short-courses for faculty and young researchers on using advanced tools and methods for informatics and big data analytics.
- Create a speaker series of high profile external speakers to both engage and inform the campus of outside thinking and activities.

The above recommended actions are expected to: (1) provide a central point for the collection of priority research informatics policy, infrastructure, and service issues; (2) foster shared standards; (3) leverage existing strengths in imaging and wireless technology innovation; (4) allow faculty an opportunity to define issues pertinent to their own research; and (5) expose students and faculty to cutting edge informatics and technology advances emanating from the community and corporate world.

Governance

To ensure cross campus orchestration, one key action item is the creation of a RISP Board. The Vice Chancellor for Research and Vice Provost will charge the RISP Board to:

- Resolve the findings in this plan into an actionable set of priorities, recommendations and actions for the entire campus and ensure that the stakeholders in informatics work together.
- Provide recommendations on mechanics for the spectrum of entities on campus to interface effectively.
- Identify informatics areas that are transformative and can capture the imagination of private and public donors and corporate partners.
Identify, investigate and propose campus policies and practices that are currently unaddressed through existing campus structures, e.g. data sharing plans, practices around meta-data, etc.

- Validate, plan and support the implementation of the immediate direction-setting actions.

**Review and Approval of the RISP Plan**

All participants of the RISP process, as well as the Deans’ Council, Academic Senate leadership, and ITPB, will be given an opportunity to review the strategic plan for comments. The RISP report will then be finalized and circulated to the entire campus for a final review.
II. Context

The Information Technology Planning Task Force (ITPTF) drafted a ten-year UCLA IT Strategic Plan, referred to as IT2020, in order to help transform UCLA into a leading 21st century public research university. The IT2020 plan represents a broad and bold general campus information technology strategic plan, within which are numerous references to the key importance of informatics to the research mission of the university. There was strong recognition that research informatics is a key enabler and differentiator for UCLA Research. The Office of Information Technology (OIT) and the Office of the Vice Chancellor of Research (OVCR) felt a critical need to further define, develop, strategize, and plan around the informatics components essential for its students, staff, and faculty in the pursuit of innovative and differentiating research. Therefore, a strategic planning process was initiated to develop a 10-year Research Informatics Strategic Plan (RISP) that focused specifically on the coordination, development and investment to build campus capacity and capability now and for the future. Accordingly, the RISP Initiative complements and augments the current IT2020 plan.

For the purposes of this plan, research informatics is defined not only as all of the data collection, generation, computing, software, storage, analysis, visualization, access, and sharing in support of research but also the supporting infrastructure, technology, tools, services, policies and governance. Like many leading academic institutions, UCLA has many innovative centers of excellence doing research on and with research informatics. There are also many areas that could benefit from investment and further development. Often, key campus investments are applied in an uncoordinated and unleveraged way, resulting from barriers to campus-wide communication, the difficulty in finding researchers with common interests or useful resources and the lack of time or resources to develop a sufficient foundation for a common research informatics vision. Moreover, opportunities and valuable resources in one area of the campus are not broadly recognized nor made available to the broader UCLA community.

The timing of the RISP effort coincides with several important events and trends: (a) UCLA is the recipient of a number of large, research informatics oriented awards in the past three years; (b) the U.S. has focused national attention and investment on Healthcare, Manufacturing, and Energy; (c) all federal agencies have organized programs around informatics and “big” data; and (d) there is a national emphasis, led by the National Institutes of Health and the National Science Foundation, to ensure that research data is accessible and shared with the broader research community.

To ensure a tangible action oriented strategic plan, this planning process focused on three key areas: (1) defining research informatics barriers and obstacles that currently exist for faculty; (2) identifying a vision of themes that, if pursued, would enable UCLA to differentiate itself as a leader in research informatics in the next 5-10 years; and (3) proposing a set of actions, policies, and organizational processes that would be responsive to both Phase One and Two. This focused approach to the strategic planning process reflects that it is neither all-encompassing nor static, but rather a tool to initiate a campus-wide dialogue that must be continued and iterative in order to be successful and transformative.
Most leading academic institutions and medical centers have made, are, or will be making a significant investment in information technology. Similarly, UCLA over the last 5 years has begun making more investments in information technology, and has also managed to remain competitive by increasing efficiency in its investments (i.e. by developing a shared cluster and data storage site).

To remain competitive the following actions need to occur from a resource utilization perspective:

- Current and future investments in research informatics should be viewed in the broader context of a proposed strategic plan to ensure they are well leveraged and impact the broadest campus constituency possible.
- A strong emphasis on institutional investments should be placed on team science and diverse cross-campus collaborations to ensure that UCLA is in a position to rapidly compete when infrastructure and research grant opportunities are identified that can substantially advance research informatics on campus.
- The University needs to accelerate its governance, investment and decision-making processes in response to the increased pace and competitiveness of research and development among leading academic institutions.
- New funding sources, in addition to grants, need to be cultivated by capturing the imaginations of donors and industry partners in applying research informatics to approaching grand challenges.
III. **Scope and Approach**

The specific goals of the RISP Initiative are to:

1. Develop a vision for UCLA and the strategic principles that will guide planning;
2. Empower and Empanel a RISP Board that will be tasked to:
   a. Develop a 5-year detailed roadmap that defines priorities and critical paths;
   b. Scope effort and approximate cost of a 5-year initiative;
   c. Develop the 1-year project and implementation plan;

Planning was divided into three phases. Phase One (Develop Core Aims for Research Informatics) was focused on building an aware campus and collecting data on the core needs and aims for research informatics. Phase Two (Develop Vision and Strategic Themes) was focused on the development of a vision and core themes for the future through individual, cross-campus discussions until a stable campus pattern of themes was established. Phase Three encompasses forming, vetting, validating and documenting the data and findings into an actionable strategic plan that includes guiding principles, priorities for implementation, five-year cost estimates and an implementation plan for year one.

The outcome of the RISP Initiative will depend on critical mass acceptance resulting from an efficiently conducted review and iteration with the campus. The review will include campus IT Governance entities, the Deans’ Council, the Academic Senate and other relevant oversight committees. With respect to the broader community, vetting activities expected or in process include:

- Conducting open campus workshops with thought leaders;
- Discussing RISP concepts with an external advisory board;
- Using a live Wiki to gather additional insight from campus; and
- Hosting a contest for students and staff, requesting cutting-edge ideas in research with prizes awarded to most promising ideas.

To facilitate the identification of core research informatics needs in Phase One, individual interviews were held with nine faculty members and five cross-disciplinary planning committees comprised of 71 faculty and staff. Each of the five committees assembled three times between December 2011 and April 2012. The five committees were:

1. Libraries; Theatre, Film & Television; Arts & Architecture
2. Health Sciences: Medicine, Dentistry, Nursing and Public Health
3. Physical Sciences, Life Sciences and Engineering
4. Management, Law and Public Affairs
5. Social Sciences, Humanities, Education and Information Science

Full membership of each committee and their chairs are provided in Appendix A.

Results from the committee deliberations were validated with a survey to all committee members as a mechanism to confirm opportunities and challenges identified during the Phase One planning sessions. From May – November 2012, and following the analysis of survey results, over 60 faculty members were interviewed and several small group meetings were held with the Institute for Digital Research and Education (IDRE); the IDRE Board; the Humanities; Arts, Social and Information Sciences Research Group (HASIS); the Clinical Translation Science
Institute (CTSI) Board; and the Social Sciences group, to identify the “big ideas” expected to influence research informatics in the next ten years. Ten strategic themes emerged from these meetings and interviews. These themes reflect a synthesis of findings from the meetings, interviews and the survey. A list of all interviewees is included in Appendix B, and a detailed write-up of each theme is provided in Appendix I.

Figure 1 below illustrates the number of faculty and thought leaders (from across campus) involved in Phases One and Two of this initiative.

The RISP process was open to the entire campus through the Research Informatics Visioning Contest. The contest invited all UCLA graduate, professional and undergraduate students, staff and faculty members to share their most innovative ideas for using data to achieve new capabilities in research, and for collecting, processing, analyzing, sharing, and leveraging research data within the campus community and beyond. The RISP committee received nearly 80 submissions and put their favorite ideas to a campus vote. The four winners each received an iPad mini. This contest ensured that no voice went unheard and allowed the RISP team to further test the emerging RISP Themes.

Throughout this planning process, the Vice Chairs of Research, the Vice Chancellor’s Cabinet, and the Executive Committee of the Academic Senate have reviewed and contributed their insight into this planning document. In addition, an External Advisory Board will review the draft plan. Members of the External Advisory Board include:

6. Michael Edsen, Director of Web and New Media Strategy at the Smithsonian Institution
7. Rayid Ghani, M.A., Chief Data Scientist for the CI's Urban Center for Computation and Data at the University of Chicago; Chief Scientist for the Obama for America 2012 Election Campaign; Co-founder of Edgeflip
8. Stan Ahalt, Ph.D., Director of RENCI (Renaissance Computing Institute), serving University of North Carolina Chapel Hill, Duke University and North Carolina State University
9. Shawn Murphy, MD, Director of the Partners Research Patient Data Registry, Harvard
10. Michael Becich, MD, PhD, University of Pittsburgh Chair and Professor Department of Biomedical Informatics
11. Daniel Masys, MD, Biomedical and Health Informatics, University of Washington
IV. Summary of Phase One Findings

Through their deliberations, the cross-disciplinary planning committees identified a series of foundation-level informatics services and support needs for researchers. All disciplines cited the need for expertise in data design and collection, analysis and interpretation, manipulation, reporting and presentation. Others highlighted support requirements such as expertise in web scraping, programming, and statistical modeling. Issues related to data storage, preservation and extraction were also common themes across most disciplines. A reliable network infrastructure with bandwidth and speed to process big data, the ability to access and manipulate data located on and off campus, and expansive space for storage (both backup and archival) was also discussed. Questions about standards and policies related to data were mentioned, including the following:

- How will research that uses non-traditional datasets and publishing mechanisms be accounted for in the promotion and tenure award process?
- How will standards and protocols be established for various data types? What are these standards and protocols?
- What are quality assurance policies and metrics for data?
- How will intellectual property and copyright issues be addressed? What are policies for new sets of data that include multiple copyright and intellectual property sources?
- What are security policies for data and how are these enforced?

In summary, these needs are categorized into these four areas:

1) **Support Services.** These are common services needed by informatics related research teams that are not specific to a singular discipline. Services include grant support, IRB advisory support, researcher training, research support, legal support, regulatory compliance and library support.

2) **Technical Infrastructure.** This includes the need for a campus-wide network infrastructure that is secure and supports high performance computing, data storage, file transfer, and universal access for all UCLA constituents.

3) **Policy, Organization and Governance.** Decisions and organizational structures are needed in the areas of data storage and archival standards, storage policies, intellectual property, faculty incentives, roles and responsibilities for access, etc, are necessary to best support informatics related research at UCLA.

1) **Tools.** Tools to support researchers include universal access to software, templates, processes and other technology-based tools.

To identify the priority opportunities and challenges associated with the four categories of basic needs, a survey was sent to all members of the five committees during Phase One. The data in Figure 2 represents the top three opportunities and challenges segmented by each committee.

---

1 Detailed examples of the needs and services in each category are provided in Appendix C.
### Figure 2: Priority Opportunities and Challenges identified by the Five Committees

<table>
<thead>
<tr>
<th>Opportunities/Challenges</th>
<th>Libraries, Theatre, Film &amp; Television, Arts &amp; Architecture</th>
<th>Health Sciences: Medicine, Nursing Dentistry, Nursing Public Health</th>
<th>Physical Sciences, Life Sciences &amp; Engineering</th>
<th>Management, Law &amp; Public Affairs</th>
<th>Social Sciences: Humanities, Education &amp; Info. Sciences</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Opportunities</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provide access to tools to mine and manipulate data</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Create a data repository that allows for easy storage, linkages, and transposing data over time</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Provide access to commonly used databases</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provide local access to super computing resources (e.g. Extreme Science and Engineering and Discovery Environment- XSEDE)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Develop a campus-wide inventory of all software licenses</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Negotiate data licenses centrally to eliminate redundancy across campus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Create a data repository for sharing data sets across disciplines</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Provide training in data management and preservation techniques</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Provide computer programming support services</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Develop a campus-wide inventory of all software licenses</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Provide access to commonly used databases</td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td><strong>Challenges</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Local campus expertise in research informatics expertise is not widely known and is not easily accessible</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lack of standards for data archival (type of data, duration, access, ownership, management, etc.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Inadequate local storage space for research</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

Figure 2 continued on next page.
There are no standard protocols for cataloging and storing code or solutions developed by individual researchers or graduate students. When these individuals leave the university the information is lost.

Data is lost or becomes inaccessible when storage formats change and/or technology is upgraded because the data can no longer be extracted.

There are no formal funding mechanisms to support research equipment and staff between grant-funded projects.

Data Backup systems are not standardized across campus.

There are no formalized mechanisms to facilitate collaboration and data sharing between researchers.

There is no central resource that inventories all database subscriptions.

Key vision themes also emerged from the cross-committee deliberations. These included the desire for:

- A transparent, inclusive environment that supports interdisciplinary research using cutting-edge informatics technologies and methodologies, which will enable the researchers to acquire, analyze and share data with UCLA colleagues and external constituents, including the general public on a global scale.
- A best-in-class infrastructure and services model that is simple, reliable and flexible, which will meet the diverse informatics needs of researchers.
- Innovative informatics technologies in support of research to be developed and funded, to position UCLA as a model, premier global research institution.

**Action Planning for Basic Needs and Services**

Subsequent to the five committees, identification of certain basic needs and services by OVCR, OIT, IDRE and the Library, came together to discuss an action plan for addressing these needs. The needs were divided into four categories, based upon the immediacy in which they could be addressed. Implementation of many of these needs and services began in November 2012. For the status of implementation, refer to Appendix D.
The Four Categories of Needs:

- **Addressed Immediately.** The first category included those needs that could be addressed instantly through a newly developed Research Resource Portal (See Figure 3). The Research Portal identifies, describes, and provides links to various services and tools that are available to researchers both on and off campus. The needs are organized under broad topics, such as Data Collecting, Sharing and Storage; Project Management and Efficiency Tools; and Research Related Policy, Regulations and Forms. The portal is important for aggregating, organizing, and structuring a wide spectrum of information, tools and people resources. Equally important is the usability and the ability for researchers to find, understand and access resources at the point and context of need. In developing researcher usability and accessibility, not only was the Portal structured, but a process was also established for maintaining resource links and descriptions that contextualize the resources with respect to capability and scope. These resources can be found using the search function. In addition, the campus knowledge base has been established as a Q-&-A front end in which researchers can ask The Knowledge Base questions, get answers from the community and link immediately to the relevant resources. Lastly, the portal is being included in faculty and graduate student community “FacTech” programs and orientations as well as being coordinated with department help desks.

- **Service and Support That Already Exist.** The second category includes services and support that already exist on campus, but which are now being coordinated and expanded to respond to basic RISP needs.

- **Under Development.** The third category describes upcoming services and tools being created to address needs that have not yet been met by the University, but involve ongoing planning by specific groups on campus, such as the Library and OIT.

- **Require Discussion.** The fourth and fifth categories concern needs and policy issues that cannot be immediately addressed because they relate to larger themes and policy issues, and require cross-campus strategic planning. These needs and policy issues will be

---

2 An IT group aimed at enhancing IT literacy and innovation among UCLA Faculty and Graduate Students.
charged to a RISP governing committee. For more information see Sections V-VII. Some examples of the basic needs that fall within each of the above five categories are provided in the table below.

### Examples of Needs and Policy Issues from Phase One

<table>
<thead>
<tr>
<th>Knowledge Management via the Research Resource Portal</th>
<th>Ongoing Services and Support</th>
<th>Upcoming Services and Tools</th>
<th>Basic Needs Linked to Themes</th>
<th>Basic Needs and Policy Issues Linked to Organizational Recommendations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training and access to statistical software</td>
<td>The Knowledge Base</td>
<td>Data Registry hosted by the Library</td>
<td>Support and mechanisms to facilitate collaboration and data sharing</td>
<td>What emphasis should be on homegrown analysis tools</td>
</tr>
<tr>
<td>Training and access for collaboration tools</td>
<td></td>
<td>On-campus cloud storage</td>
<td>System for identifying and pooling researchers or other FTEs to share hiring costs</td>
<td>Policy for making decisions about data stored locally versus off site</td>
</tr>
<tr>
<td>Assistance in designing, deploying &amp; analyzing surveys</td>
<td></td>
<td>Thesis and dissertation services and support</td>
<td>Simplify file transfer process and identify management</td>
<td>Standards for granting permission to access data, as to alumni, the public and those internal to UCLA</td>
</tr>
<tr>
<td>Access to workflow tools</td>
<td></td>
<td>Boilerplate templates for various documents, such as budget justifications, data sharing and repository plans, and biosketches</td>
<td>Infrastructure that enables new models of research using large data; addressing data storage issues for large data</td>
<td>Policy for storing digitized data</td>
</tr>
</tbody>
</table>
V. Phase Two: Summary of Strategic Vision Themes

Ten major themes emerged during Phase Two of the study. As illustrated in Figure 5 to the right, Themes 1 and 9 address the types of researchers that will be engaged in informatics. Themes 2 and 3 address the types of tools, support services and infrastructure necessary to enable collaboration amongst researchers. Finally, Themes 4 through 8 and 10 address the various types and formats of data that will be collected, stored, and analyzed in the growing field of informatics. Detailed descriptions of each Theme and the associated research challenges are summarized in Appendix H.

Ten Themes

1. Cross-disciplinary Collaboration

Cross-disciplinary collaboration and partnerships will increase opportunities to share ideas and think about complex research problems together. Innovative approaches for generating, collecting, and analyzing data to bridge disciplinary languages, dictionaries, and areas of interest will provide vast opportunities for cross-disciplinary researchers to share ideas and think about complex research and global problems together.

Implications

- Cross-disciplinary approaches will allow the institution to capture a broad range of contextual lenses for data analysis and to encourage innovative data-driven applications.
- Cross-discipline collaboration will serve as a catalyst for invention that traverses campus boundaries and advances and strengthens the potential to develop global solutions to the world's toughest problems through research discoveries.

Key Examples

- **The National Ecological Observatory Network (NEON).** NEON is a continental-scale ecological observation system for examining how the biosphere is changing in response to human activities. The network is a global collaboration of researchers and scientists across diverse academic and scientific institutions. NEON uses an open access approach that draws input from scientists and researchers, and will provide data and information products to scientists, educators, planners, students and the general public.
- Several universities have already established collaboratories where researchers from different disciplines come together to solve big problems, including:
a. University of Michigan’s Institute of Social Research. This is the largest academic social science survey and research organization in the world. Their annual budget is $80 million and supports the research of more than 250 scientists from 20 disciplines, including psychology, business, economics, public health, demography, statistics, and engineering.

b. MIT’s Media Lab. With an annual budget of $35 million a, MITS Media Lab applies an “atelier-style” approach to collaboration across 350 projects unconstrained by discipline type. Researchers focus on innovation in technologies that transform basic notions of human capabilities.

c. Harvard’s metaLAB. MetaLAB is a teaching and research hub that supports exploration of questions about the human experience in a connected world. This institute facilitates collaboration opportunities across academia, industry, and the public sphere.

d. UCLA’s Institute of Environmental Sustainability (IOES). IOES enables interdisciplinary collaboration across disciplines to develop novel environment and sustainability solutions and initiatives.

2. rEcosystem (Shared resources)

An rEcosystem provides institutional informatics connectivity, key shared resources that extend capacity and the tools and capability, all which form the institutional “nervous system” and the rich functionality that includes distributed resources, capabilities and expertise brought into one coherent capability; the links through which data can be moved and methodologies accessed; and institutionally leveraged infrastructure that builds individual capability far beyond that which can be developed locally.

Implications

- The rEcosystem will significantly speed up the researcher tasks of discovering resources and orchestrating their use to do integrated research.
- With respect to facilities and technical foundations, the rEcosystem will connect and integrate an end-to-end system of resources that are workflow-driven and data-centric, while simultaneously solving challenges in network, computation, data visibility, mobility, access and transport, and storage. This will be done as one cyber physical system (following the definition by the National Institute of Standards and Technology – NIST), not as separate components.
- AT UCLA, the rEcosystem will integrate data centers across four buildings (Math Sciences, California NanoSystems Institute (CNSI), the Gonda Science and Genetics Research Center, as well as the Terasaki Life Sciences Building). The infrastructure design allows other buildings on campus to be easily connected, such as the Young Research Library, which houses modeling labs for the Digital Humanities at UCLA and has large data repositories in the Social Sciences and Theatre, Film and Television, as well as the library which has a large collection of archived research data.
- The installation of an optimized 10/100Gbps research network by itself is an essential but insufficient step to properly support use by campus-wide scientific applications. The workflow design and execution technologies will also be needed to link data resources, HPC, web services, software tool libraries, and other technologies. Once a system is created that incorporates the above two elements, there will be an increased use of data processing.
workflows, which will require a greater level of network bandwidth. This increased data processing workflow will be drawn from large-scale science databases and leveraged through web-based tools and services. It will utilize HPC infrastructure and cloud storage.

- The rEcosystem is fundamentally a shared institutional investment to build a shared resource. It requires institutional and local buy-in and investment such that there will always be value derived beyond that which can be generated by individual researchers.

**Key Examples**

- **Genetic Bioinformatics**: Current single nucleotide polymorphism (SNP) technologies encode gene level expression on standard chips of 1-2 million SNPs (e.g. Affymetrix 1M). When obtained from a large sample of study participants, gene expression levels from alleles of these genes can be quantified and used to identify correspondences between genetic and phenotypic variation. Such technology has been applied to genome-wide association studies (GWAS) of the genetic components of many human traits and diseases. However, recent technologies now permit characterization at the level of the “exome” - the complete description of transcribed genes present in an organism or individual. These new methods can be expected to completely eclipse the amount of genetic information available for GWAS-type studies involving human, as well as non-human, species. rEcosystem network capacity will be essential for moving exome-level information from the gene-sequencing units located in the Gonda Building, for instance, to data storage systems located in the UCLA Cloud, or for analysis using the Hoffman2 Shared Cluster. Workflow technologies will be needed to schedule file movement over the network, as well as to address computational challenges such as sequence alignment, partitioning, and subsequent informatics. Coupled with other large data sets (e.g. neuroimaging data, above) the combined processing of genetic and physiological data will require high-bandwidth networking and efficient workflow cyber infrastructure to encourage new discovery. The UCLA rEcosystem has been designed to specifically support such large-scale, multi-source data analyses.

- **Electron Microscope Imaging**: Recent advances have been made in cryo-electron microscopy (cryoEM), an important imaging tool for major applications in both medicine and nano-biological research. Researchers can use cryoEM to visualize a broad range of assemblies or nanometer-scale structures at near-atomic resolution and in three dimensions. This imaging method covers a scale range from tens of micrometers to angstroms and provides valuable structural information for numerous scientific disciplines including cell biology and microbiology, as well as medical, biomolecular, molecular, pharmaceutical and materials sciences. Storage of such electron and spectrometry data using the UC Cloud would allow data to be readily shared with computer science collaborators applying novel processing algorithms and data visualization techniques. The UCLA rEcosystem 10/100Gbps network connectivity to virtualized storage, database interoperability, and processing workflow technologies will greatly facilitate and streamline these cross-disciplinary activities.

3. **Enabling a broader base of researchers**

   *Easier-to-use, self-guided and more highly abstracted transformative tools and services that reflect informatics expertise will enable a broader base of researchers to conduct novel database research without having to develop or invest in the same expertise.*
Implications

- Strategies and tools that facilitate overcoming barriers of knowledge, personnel resources, infrastructure, and IT solutions that provide even a small incremental improvement in the incorporation of information technology will have enormous impact given the number of researchers that have these challenges.

- New models for research informatics support such as collaboratories, will be established, to support researchers who may be in silos or that do not have the resources to establish independent infrastructures and support systems. Access to standardized approaches to data documentation and to definitions will be accessible to all faculty, including those in fields that have been under-developed.

- A campus marketplace for research informatics faculty and staff will be created by fostering new pathways to identify and find collaborators with competency and skills in data collection, management, and analysis.

- Broader access to technology and informatics will result in new and increased research opportunities and discoveries in a more efficient manner. New tools for data management and analysis will need to be developed and streamlined into well-documented workbenches to allow less sophisticated users a lower barrier for entry in utilizing research informatics. Data collection and analytics will become more efficient across all research.

Key Examples

- **IDRE-HASIS core** is a catalyst for advancing and aligning south campus scientists that have deep computational expertise with the social and cultural interests of north campus disciplines such as humanities and social sciences. Increasingly, new digital tools, methodologies and technologies (e.g. advanced searching, data mining, automatic metadata generation, visualization of cultural and social data, real-time interactivity, etc.) are shaping research in north campus disciplines. IDRE enables the development of cross-campus collaborations and relationships that were traditionally impaired by geographic constraints.

- Researchers across universities and institutes are using the panel, **Rand Corporation’s American Life Internet Panel** as a way to access research data without having to replicate the underlying infrastructure. For a small fee, researchers submit research questions to be integrated into a twice-monthly survey tool with participation from over 5,000 individuals, ages 18 and up. The survey results (research data) are made publically available after an initial period of time.

- The Pittsburg Computing Center has developed an open source workflow tool called **Galaxy** that supports biomedical researchers in performing, reproducing and sharing complete analysis of genome data. Professor Matteo Pelligrini, UCLA Department of Molecular Cell and Developmental Biology, is customizing this tool to avoid replicating advanced workbench tools.
4. **Data Ownership and Big Data**

Big data has three attributes: volume, variety and velocity. Volume represents the scale of data; variety represents data’s many forms, such as structured, unstructured, text, and multimedia; and velocity represents dynamic, real time data will be created by the ability to more readily access and collect data that extends beyond the walls of the institution, and the ability to store and analyze large amounts of disparate data (or big data) that investigators may not wholly generate or own in their own research. Accelerated and novel opportunities for non-traditional research analysis and decision-making.

**Implications**

- Data accessed from tools such as publicly available government databases, social media, external researchers and research organizations will be synthesized with data sets that originate from campus researchers. Connecting a variety of data from disparate external sources will inform new ways of synthesizing and storing data.
- Large and real-time data accessibility across all disciplines will be expansive and enable researchers to ask novel questions and extract unique insights that have not been possible with smaller subsets of data.
- Bi-directional flow of information will create an expansive community where rules for sharing and annotating data responsibly become important principles of being a member in good standing.

**Key Examples**

- **The UCLA Communication Studies Archive (CSA).** The CSA is one of the largest archives of news and public affairs television programming in the world, storing video from the late 1970s to the present. Cutting edge opportunities exist in developing analytical, text-parsing tools to better analyze this set of big data for example, by graphically subdividing data and recognizing commercials and advertisements.
- **UCLA Professor Matteo Pelligrini’s lab.** Professor Pelligrini’s lab in the Department of Molecular Cell and Developmental Biology is developing computational methodologies to interpret and annotate genomic data. The lab is also developing suites of tools for DNA methylation data analysis and tools to annotate genomes using RNA sequences.
- **The Billion Prices Project at MIT.** This is an academic initiative that collects prices from hundreds of online retailers around the world on a daily basis to conduct economic research. One part of this initiative, the US Daily Index, is designed to provide real-time information on major inflation trends based on an average of individual price changes across multiple categories and retailers. These statistics are available for public use.

5. **Real-time Dynamic Data**

Real-time dynamic data and analysis will transform traditional research approaches and methodologies by accelerating the generate-analyze-apply-learn research cycle.
Implications

- Real-time data analysis will transform into a sophisticated practice of generating and applying data-driven “intelligence” throughout the lifecycles of design, engineering, planning and production.
- Systems will use networked, information-based technologies to integrate intelligence in real-time across an entire enterprise and will use data-driven modeling, simulations, and Key Performance Indicators (“KPIs”) to communicate optimal actions in real-time.
- All needed information will be available when it is needed, where it is needed and in the form that is most useful.
- Advances in the use of mobile devices and applications and the expansive use of sensors will accelerate the generate-analyze-apply-learn (a rapid, dynamic, and iterative) research approach.

Key Examples

- **The Wireless Health Institute.** The Institute conducts cross-disciplinary research actively using real-time dynamic data in research. The Institute received a $10M grant to fund pilot projects that progress to clinical trials with the goal of becoming large-scale research projects. One focus of research includes continuous monitoring of the human body to innovate in the field of personalized and customized medicine. Currently there are 50-1,500 subjects that generate large quantities of raw data, collected, analyzed and studied in real time.
- **Arab Spring and Twitter Feed.** A team of faculty members from UCLA’s Department of Humanities was able to instantaneously collect and display Twitter feeds from Egypt onto Google Maps during the Arab Spring, so that researchers, analysts and the general public could see and follow where the activity was happening in real time. Researchers at UCLA tracked and analyzed the social media state of the revolution in real time building insights as they occurred.
- **Real Time Data and Radioactivity Levels.** UCLA researchers and staff established real time databases first to track immediate resources available to people in the wake of the Tsunami that hit Japan and later established real-time data on the radioactivity levels of areas surrounding the nuclear reactor exposure area.
- **Arsenic Levels in Rice Paddies.** UCLA Investigators from Engineering, Computer Sciences and Statistics were able to discover the cause of toxic arsenic levels in the rice paddies of Bangladesh by studying the pH levels in the groundwater as it was being collected in real time by sensory equipment.

6. **Multi-Use Data**

> An increase in multi-use data will blur the boundaries of research, business, and operations, allowing research to more seamlessly integrate into business workflow and operations.

Implications

- Innovation in research methods will inform a cultural paradigm shift that will also transform business processes to account for more open and shared access of data.
- Policy and governance in organizations with cross-purpose data will consult with local IRBs and consider other compliance standards in adapting to managing data that is now multi-use.
Considerations in shifts in business operations will focus on how business and research are supported simultaneously.

- Decision and action workflows will leverage new systems to accommodate multi-user/cross-purpose access to data.
- Consolidating data resources into multi-use data sets will help streamline projects and processes and improve economic efficiency in research, business and operations.
- Taking into account the requirements of dual-use or multi-use data will allow better collaborations with industries that use learning cycles quicker and brisker than traditional academics.

Key Examples

- **The UCLA Graduate School of Education.** The Graduate School of Education frequently creates and publishes cutting edge learning, testing and assessment tools for primary schools that use technological platforms like social media applications in place of standard textbooks. These tools are often experimental and simultaneously studied by UCLA researchers as they are being used in the classrooms.
- **The UCLA Department of Political Science.** In 2006, Lynn Vavreck, from the UCLA Department of Political Science, along with 100 researchers from 38 other universities, ran the first nationwide presidential poll conducted exclusively on the Web. This study, which polled 39,000 voters (more than any other election poll ever taken) not only served to predict election results, but also provided new opportunities for inquiry to a broad range of researchers at a rapid rate.
- **The UCLA Medical Center.** De-identified clinical data from UCLA’s medical center can be used for research in various departments across campus including the liberal arts, humanities and social sciences. For example, an economist may combine de-identified patient data with housing market data from Zillow.com to draw a correlation between health risks and standard of living.

7. **Image Data**

Images are data. Imaging is the capture, manipulation, storage, and visual representation of data. Imaging represents the intersection of technology and data collection, management, and analysis with implications relating to institutional infrastructure. Image data research will continue to grow in sophistication as the analysis of image features increases in granularity and descriptive detail.

Implications

- The role of imaging in research will continue to grow more complex to advance exploration, explanation and communication of data. Key commonalities include, quantization, resolution, dimension reduction, feature detection, distortion, perception, noise, compression, presentation, sensitivity, and uniformity.
- Fields such as physics, chemistry, medical imaging, astrophysics, nanoscience, mathematics, design media arts and statistics will drive cross-campus partnerships to support greater access to the use of imaging to facilitate transformative research outcomes.
- The decisions made as an image moves from its raw state to the final processed state will be influential in research outcomes. What is lost in the evolution from the raw to the final processed state? Does the data, during its transition, inform novel outcomes?
• The campus will develop a principled approach to manage the collection, representation, and analysis of images over a wide range of dimensionality.
• Groundbreaking discoveries will result in specific focus on (a) analysis and modeling of visual patterns and images, (b) visual perception and use of visual information in decision-making, (c) temporal images and functional imaging.
• Advanced methods of extracting and utilizing information in images with a focus on data compression and compressed sensing, will likewise contribute to new discoveries.

Key Examples
• \textbf{Digital Roman Forum}. The UCLA Cultural Virtual Reality Laboratory (CVR Lab) has created a digital model of the Roman Forum as it appeared in late antiquity. Among other things, the purpose of this site is to provide a virtual setting for scholarly study where history and cultural context can be readily understood and questioned through an avatar.
• \textbf{Michael Darby, Anderson School of Management}. Michael Darby is interested in creating an algorithm that converts audiovisual material from sources like Youtube.com into data that can be analyzed in a digital database. With these types of capabilities, novel hypothesis can be drawn from data that is deconstructed into smaller bytes.
• \textbf{Christopher Contag, Stanford Medical School Microbiology}. Professor Christopher Contag has created a new technology that constructs 3D images hundreds of cells deep, allowing doctors to make informed cancer diagnosis on the spot rather than sending samples to pathology labs and waiting for results to come back.

8. \textit{Citizen Scholar}

\textit{The emergence of the citizen scholar, non-traditional research teams, and the broader community will increase demand, access and contribution to shared data. It is becoming more common for individuals that are not employed by a college, university or other research-based organization, to engage in research purely as an enthusiast, as a "citizen scholar," which in time will increase demand, access and contributions to shared data sets.}

Implications
• Access to data and the products of data analysis will allow broader interpretation of results and outcomes of research projects, and potentially inspire new research areas.
• Sharing the same principles, approach and philosophy about research, the citizen scholar will serve as a bridge between the traditional researcher and the broader community. This bridge will expand institutional intellect, creative insight and capacity for addressing complex problems.
• Resolving the issues around data documentation, access, sharing and collaborative tools required for a citizen scholar to truly participate in research will ensure all the other communities also benefit.
• The citizen scholar will be an important partner to drive transformation in the UCLA research community through both analysis and contribution of significant intellectual capital to diverse research areas. The armchair astronomer, for example, now has more access than ever to materials on par with academics, which fuels their ability to contribute intellectual power to their discipline.
• UCLA will leverage the multi-cultural and diverse strengths of the community and its individuals in innovative research that is "community-centered" and/or "patient-centered."
Key Examples

- **Hypercities.** Created by Professor Todd Presner, Hypercities is a digital research and educational platform that allows any individual to use interactive tools to explore and contribute content to the social, cultural and political history of a community over time. Content is being developed in this global platform for cities such as Los Angeles, New York, Chicago, Lima, Berlin, Tel Aviv, Tehran, Saigon, Tokyo, and Shanghai, with other locations soon to follow. Individuals can create content such as community narratives in video, or add photos or other media to collections that exists along with scholarly produced research data.

- **The e-Patient Scholars Working Group.** Founded by the late Tom Ferguson, MD, the working group is leading the transformation of the doctor-patient relationship into a partnership. Patients with common conditions are using the Internet to contribute information that is universally available. This allows patients to become more proactive in managing and learning from their own health in partnership with their physicians. UCLA could become the national model of patient satisfaction turnaround.

- **Choose your own adventure videos.** Faculty within the School of Theater, Film and Television foresee a future trend in film in which viewers select divergent storylines using branching techniques in order to decide what happens next in the story they are watching. Movies are no longer static but require audience participation, and data relating to the viewers’ decision patterns could open new lines of research.

9. **Data Visualization**

Data visualization is the subset of visualization that involves the graphic display of data too complex for manual processing, and where, accordingly, the resultant imagery is typically the end result of an algorithmic process or generated from large-scale data sets. Data visualization encompasses a broad range of analytic tools and techniques that include statistical visualizations, GIS and 3D modeling all which share the common goal of organizing data into a coherent visual display that can be readily interpreted and understood.

Implications:

- The intricacies of research outcomes that use complex data from multiple sources will necessitate a focus on strategies for how data is presented, interpreted and consumed in human centered settings. Researchers routinely consider how to present information in a succinct manner and design systems based on a particular population, situation or use of the data.

- Data visualization is also important to streamlining decision-making, i.e. clinical environments. Visualization will therefore not only be an interface but will also be an integral component to the outcomes or conclusions from the data.

- Data is often displayed in the form of movies, which are especially revealing when used to display scientific simulation data that includes multiple time steps. Movies will also allow researchers to rotate or transform objects.

Key Examples:

- Typical examples of visualization include laser, fluid, and climate simulations, network analysis and geographical mapping of populations. Statistical modeling is widely used for longitudinal analyses, factor models, classifications, interaction models, probability models
and marginal effect models to name a few. 3D computer modeling involves the generation of geometry either manually, from data, or from a scanning process.

- **UCLA Visualization Lab.** Dr. Alan Garfinkel and his team of researchers at UCLA’s Department of Medicine have been researching heart fibrillation and drugs that can stabilize the heart before total cardiac arrest occurs. In a healthy heart, waves of electricity make the muscle contract and produce a rhythmic heartbeat. In an abnormal heart, those electrical waves become unstable and fragmented, which produces chaotic mini-contractions instead of a steady rhythm. Called ventricular fibrillation, these episodes are the leading cause of sudden cardiac death. Using UCLA's Visualization Lab it is possible to see a visual representation of heart fibrillation and the effects of particular drugs on cardiac arrhythmias. To simulate one second of heart activity in one section of the heart required 7 trillion calculations run over a period of several days on a cluster of high performance computers.

- **Visualizing Atmospheric Data.** Atmospheric Sciences Professor Bjorn Stevens, Research Scholar David Beaudry, a clarinetist and sound designer who received a Doctorate in clarinet performance from UCLA and Portal Development Coordinator, Joan Slottow, joined to visualize and sonify atmospheric data. Professor Beaudry used his expertise in sound design and music to help Professor Stevens understand and interpret cloud data produced by simulations of non-precipitating cumulus clouds in the tropics through sonification, the use of sound and music in the display of data. The impetus behind this project was to examine how visual artists, sound designers, and musicians can help visualize and sonify complex data sets by creating new visual and sensory approaches to help in the understanding of the data. This project specifically focused on developing more universal sonification tools to aid in the interpretation of atmospheric scientific data.

10. **Mobile and Social Networking**

Information and communication technologies have achieved unprecedented pervasiveness and ubiquity over the past decade (e.g., mobile phones, Google search, Facebook, Twitter, Pinterest, online maps). Never before has there been a time in which billions of people can interact and conduct business. The wireless revolution is producing an ever widening and thickening blanket of human centered and sensor-based data.

**Implications:**

- Mobile and social media has transformative powers, and will allow faculty to form powerful connections and reach new people that previously could not be accessed from the ivory tower. Social media platforms will inform every step of the research process, helping faculty meet other academics interested in their work; get up-to-date information on their industry; provide data for new research ideas and possibilities for innovation; provide feedback during research; and assist in the promotion of the published work.

- Advances in engineering and computer science will enable the design of powerful home and mobile technologies that can augment functional independence and daily activities of people with physical impairments, disabilities, chronic diseases and the accumulative impairments associated with aging. These home-health and mobile-health technologies will serve as monitoring devices of health and activity, feedback reinforcement for risk factor management, and outcome measures for individual care and large clinical trials.
Every discipline that has human beings as its focus now has new channels for both understanding and effecting impact: psychology, education, economics, development studies, public health, medicine, art and architecture, and law. Each of these disciplines can now, for the first time, readily obtain data around individual and aggregate behaviors and phenomena that were previously inaccessible; and each can build on the ability to communicate with ‘instrumented populous’ when proposing policy and interventions. Students, staff, researchers and educators will need the ability to quickly, affordably, interactively, and iteratively explore how to use these new media in real settings to answer real questions with real people.

Key Examples:

- **School of Engineering and Applied Science (SEAS).** Researchers in SEAS have developed Wi-Fi network technologies for high bandwidth wireless data exchanges to be used in data access applications such as police cars and transportation units.

- **Ohmage.** MobilizeLabs has developed and is making available Ohmage (http://ohmage.org), an open source mobile to web platform that records, analyzes, and visualizes data from both prompted experience samples entered by the user, as well as continuous streams of data passively collected from sensors or applications onboard the mobile device. Ohmage additionally includes rich system and user analytics to instrument the act of participation itself and ultimately to contextualize and better understand the factors affecting the quality of collected data over time. A number of community data generated studies are underway: (1) public health community exercise challenge, (2) HIV positive patients and substance abuse, (3) low nicotine cigarettes and (4) young cancer survivors.

- **Mobil Web Framework (MWF).** UCLA has deployed the MWF, a tool that enables users to quickly and easily create device-agnostic, platform-independent, and easy to deploy framework for the mobile web. Built with the content provider in mind, the MWF allows the provider to avoid detailed device-by-device planning, implementation, upgrades, and maintenance, making good on the promise of “develop once, use everywhere.” Included in the MWF is an Online Polling Tool for collecting audience/community feedback in real time.
VI. Phase Three: Creating a New Governance System

Current Research IT Organizational Structure and Governance

In recent years, the campus has recognized the importance, value and necessity in planning research informatics infrastructure and service investments as an integrated whole. Considerable progress has been made in recent years in that many larger-scale, cross unit operational decisions about research IT are managed through the Clinical Translational Science Institute (CTSI), the Institute for Digital Research and Education (IDRE) Boards and the Libraries, in conjunction with direction and policy through the Information Technology Planning Board (ITPB). Although these efforts have been tremendously important and productive in many areas (solid line silos have given way to dotted line breaks), there remains untapped opportunity, especially as the research advantages of shared expertise and capability intensify. While IDRE was formed to address these issues with computational based research, the emergence of Informatics and “big” data required that a broad survey of campus efforts in informatics be conducted. Much campus expertise and many decisions with potential campus impact remain local in schools, divisions, departments and various research entities. Broadly speaking, the RISP exercise has pointed to the need to more tightly define relationships and processes for campus investment, consensus building, collaboration, and prioritization to achieve better integration of resources and at the same time better capability for local research needs. See Figure 6 on the next page.

Proposed Research IT Organizational Structure and Governance

The proposed organizational change will initialize a formal Research Informatics Strategic Planning Board, as shown in Figure 7, to resolve the findings in this plan into an actionable set of...
priorities and actions for the entire campus. The RISP Board will be organized under the auspices of the OVCR and OIT.

Figure 7

**Informatics Planning**

**INITIAL Sponsorship, Oversight & Governance**

**Sponsorship & Oversight - Operating Decision & Outcomes**

**Governance – Direction, Investment & Policy**

**OVCR & OIT**

Research Informatics Strategic Planning Board (N~15)

**Charge to RISP Board**

The Vice Chancellor for Research and the Vice Provost for Information Technology will charge the RISP Board to:

- Resolve the findings in this plan into an actionable set of priorities, for the entire campus and ensure that the stakeholders in Informatics work together.
- Provide recommendations on mechanisms for the spectrum of entities on campus to interface effectively.
- Identify informatics areas that are transformative and can capture the imagination of private and public donors and corporate partners.
- Identify, investigate, and propose campus policies and practices that are currently unaddressed through existing campus structures, e.g. data sharing plans, practices around meta data, etc.
- Validate, plan and support the implementation of the immediate direction-setting actions (See *Taking Action*).
- Establish and hand off a review process in which the status and impact of informatics actions are measured against an annually reviewed strategic direction plan.

**Composition of RISP Board**

UCLA has strong cross-campus and interdepartmental institutes that are making substantial contributions in the area of research informatics, such as the CTSI, IDRE, and the CDH. However, there are also many School, Departmental and Divisional Research Team Centers and Units of Excellence all across campus (denoted in Figure 8 by intersecting bubbles). The RISP
Board will be composed of faculty, administrators, and students who have a deep understanding of research informatics yet a broad rather than insular campus perspective.

Members of the RISP Board will be appointed by the Vice Chancellor of Research and Vice Provost of Information Technology in consultation with relevant campus groups and will be drawn from faculty, campus administration and students. The goal will be to maintain a nimble, action-oriented RISP Board, therefore the size of the board and the frequency of meetings will be important considerations initially. The composition and activities of the RISP Board and its role in the governance process will be reviewed annually.
VII. POTENTIAL HIGH PRIORITY CAMPUS INITIATIVES

A number of recommendations emerged from the planning process and set into motion a series of directional changes for UCLA that address an institutional strategy and build institutional capacity where most useful, while also recognizing the distributed and domain-driven nature of the research. These actions support the cross-campus themes identified in the strategic planning process. These action-oriented recommendations are split into two categories: 1) Transformative Recommendations and 2) Catalytic Recommendations. The role of the RISP Board and campus impact are described for each recommendation.

Top 4 Transformative Recommendations

1. Establish and invest in a campus wide Informatics Scholar Program in which advanced undergraduate, graduate, and post-doctoral students work with domain scholars and/or computer science, computational science, statistics, and applied math faculty and/or staff on research projects across research informatics that are strategically selected. Establish and invest in an Informatics Sabbatical Program that recruits outstanding informatics scholars to spend sabbaticals at UCLA and which allows UCLA faculty to spend sabbaticals working at leading institutions in informatics.

Transformative Potential:
- Reduces the barrier for faculty whom has an interest in using computational science but limited experience or resources.
- Improves communication, collaboration, and the efficient use of resources.
- Promotes and facilitates self-organization into interdisciplinary research teams.
- Integrates education and research in the context of broadening the use of computational science.
- Ensures that UCLA will be at the forefront of research informatics.

Role of RISP Board: Establishes and reviews policies for the Informatics Scholars and Sabbatical Programs.

2. Establish workshops for non-health sciences faculty and graduate student researchers to learn and discuss health data marts available for analytical use.
Transformative Potential:
- Reduces the barrier for non-health sciences faculty who have an interest in using health sciences data.
- Facilitates the identification of new and novel uses for health data.
- Leverages one of the largest patient-centered health system data repositories in the United States

Role of RISP Board: Ensure that health and non-health informatics research is coordinated, including governance issues associated with the broad use of health science data.

3. In partnership with the IDRE and CTSI, create a shared bioinformatics resource to investigate the potential for shared experiences, expertise and tools as well as to define the network, computational, storage and software infrastructure to support it.

Transformative Potential:
- Ensures UCLA stays at the forefront of “Big” data analytics.
- Catalyzes collaboration and innovation in the areas of computational biology and health sciences leveraging two institutional strengths.
- Broadens faculty participation in the area of bioinformatics and computational biology.
- Involves the community and citizen scholar in assessing the Ethics of use of Genomics Data in Research.

Role of RISP Board: Work with the CTSI and IDRE to define key infrastructure needs required to advance the area of “Big” data analytics.

4. Provide a base level of no cost, managed data storage in combination with enacting research data sharing plans to meet a pervasive need and to incentivize faculty researchers to: (1) register their database content; (2) provide metadata and documentation associated with their data set so that it is discoverable and usable; and (3) store & back up data in one of several campus managed-data storage services that meet standards for their discipline/department and for sharing/accessibility.

Transformative Potential:
- Provides substantial incentives for faculty to identify and share existing data sources.
- Reduces the burden on faculty for managing the storage, backup, and data sharing requirements associated with research.
- Allows for a portal for collaborative research and identification of potential existing datasets
- Establishes policies regarding consideration of data generation, management, and curation for faculty promotion

Role of RISP Board: Define the minimum requirements for access to the data storage services, help define the package associated with this data storage service and help with the broad campus communication to foster participation.
Top 6 Catalyzing Recommendations

1. Institute a Faculty Technology Awareness (FacTech) Program for research informatics and take advantage of already existing FacTech department programs which function as faculty/department-sponsored ‘speed dating’ events.

   **Catalyzing Potential:** Allows faculty and departments to drive focus to informatics in research, reaching late adopters of technology.

   **Role of RISP Board:** Help identify faculty champions in each department.

2. Establish and invest in an Informatics Sabbatical Program that recruits outstanding informatics scholars to spend sabbaticals at UCLA and which allows UCLA faculty to function a ‘One-Stop-Shop’ for researcher consultation on campus resources, policies, processes, training and capacity-building, in combination with a network of resources hub.

   **Catalyzing Potential:** Facilitates easy faculty and student access to resources required to be effective in data-driven research. Also allows for a central point for the collection of priority policy, infrastructure, and service issues that the campus needs to evaluate and prioritize to resolve.

   **Role of RISP Board:** Review and evaluate issues identified by the “network of resources hub.” The Board will also develop new policies and clarify existing policies to facilitate and streamline effort associated with research informatics.

3. Instantiate an Imaging Symposium Series for UCLA museums and campus departments on image archiving and the standardization and use of image metadata.

   **Catalyzing Potential:** Fosters innovative and collaborative discussion in an area of tremendous activity on campus, while fostering shared standards.

   **Role of RISP Board:** Facilitate collaboration and work to advance the area of imaging science with industry partnerships.

4. Run Mobile Technology and Social Networking research innovation application contests or pilot grant awards where contestants/applicants develop mobile and social networking research applications that can be repurposed and used by other researchers for similar applications.

   **Catalyzing Potential:** Exposes UCLA faculty to cutting-edge ideas or technology from external academic, community, and corporate thought leaders.

   **Role of RISP Board:** Define a transdisciplinary pilot grant award or targeted technology contests, and then evaluate submissions.
5. Establish and invest in **big data analytic workshops** and summer short-courses for faculty and young researchers on using advanced tools and methods for informatics and big data analytics.

**Catalytic Potential:**
- Improving knowledge and skills of UCLA research community.
- Provides opportunities for cross-disciplinary engagement.
- Facilitates easy faculty and student access to resources required to be effective in data-driven research.

**Role of RISP Board:** Establish and review policies and assist with determining subject topics for the workshops and summer short-courses.

6. Create a **high profile informatics speaker series** to both engage and inform the campus of outside thinking and activities.

**Catalyzing Potential:** Leverages UCLA’s national presence in wireless technology by expanding the use and application of this technology in novel research areas.

**Role of RISP Board:** Identify and reach out to prospective speakers, and highlight the intersection of a topic with UCLA’s strategic plan and goals.
VIII. Next Steps

1. Review and Endorsement of the RISP Report

This draft RISP report has been distributed to and reviewed by: (1) ITPB; (2) Council on Research Committee of the Academic Senate; and (3) Deans’ Council. This updated RISP report will be distributed for feedback to the above entities plus the over 100 faculty who participated in this process, as well as the Vice Chairs of Research and Department Chairs. After comments are taken into account, the RISP will be sent to the Provost for review and then will be made open to the campus for comment. Feedback from these groups will be incorporated into a final draft prior to final presentation to the Executive Vice Chancellor and Chancellor for review.

2. Creation of a RISP Board

A RISP Board will be created following the recommendations presented in the plan. Once formed, the draft report will be submitted for their review. Over the course of several intensive retreats, the RISP Board will develop guidelines for prioritization of the strategic projects and initiatives that were identified through the strategic planning process.

3. Analysis of Scope and Cost

Based on preliminary feedback from the RISP Board, the key strategic projects and initiatives will be defined in more detail with cost estimates. This information will be provided to the RISP Board for prioritization.

4. Finalization of the RISP

The prioritization will be used to define a one-year and three-year work plan with milestones for each project and initiative. The results will be reviewed and approved by the sponsors: OVCR, OIT, School of Medicine, and the Library. These projects will also undergo review and endorsement from ITPB. Funding requirements will be reviewed by CITI. In addition, development officers and partnerships with industry sponsors will be used as a funding strategy when possible.

5. RISP as a Living Document

Given the rate of technological change and advancement in informatics, the RISP report will necessarily be a dynamic document, updated at regular intervals, with input from campus leadership and stakeholders. Additional perspective will be provided through an external advisory board.

Authors:
Jim Davis, Ph.D., Vice Chancellor for Information Technology
Arash Naeim, M.D., Ph.D, Assistant Vice Chancellor for Research
Mary Beth Baker, Consultant
CONTRIBUTORS:
Warren Mori, Ph.D., Director, Institute for Digital Research and Education
Ornah Medovoi, JD, Director of Special Projects, OVCR