Executive Summary

This decade has seen research informatics – the conduct of research involving data – rise to unprecedented capability for vastly better understandings, unimaginable new insights and new solutions at previously unthinkable scales. Data from images, observations, and sensors, past and present, can be expressed digitally together to form research databases of unprecedented scale, scope and substance. Computational power and storage has reached ‘atomic’ levels of fidelity in describing the fabric of physical, biological, behavioral, social and humanistic expression. The pervasiveness of technology in the hands of much of the world’s population has created yet-to-be realized data and its impacts. These have all occurred while the problems of global interest and their solutions have become more complex, more inter- and trans-disciplinary and with decidedly more integrative and involved partnerships.

In 2007, the UCLA Information Technology Planning Board (ITPB) recognized the growing importance and need to take action on an institutional strategy for research informatics as a key enabler for UCLA research interests and as a differentiator for UCLA research leadership. The ITPB discussions emphasized that research substance, leadership and competitiveness would depend not only on strengthened focus to enable cross-disciplinary data and analysis and the need to build interdisciplinary research teams but also the ability to build campus capacity by sharing informatics experiences and tools across disciplines and to build infrastructure that no one research group could afford to build alone. Informatics, as a strategic direction for UCLA, was carried forward into the IT2020 Strategic Plan.

With the need for an institutional strategy growing more acute and at an increasingly accelerated pace, the Office of Information Technology (OIT) and the Office of the Vice Chancellor of Research (OVCR) felt a critical need to move more rapidly and further define, develop, strategize, and plan around the institutional informatics components essential for its students, staff and faculty in the pursuit of innovative and differentiating research. A strategic planning process was initiated to develop near term actions while informed by a 10-year Research Informatics Strategic Plan (RISP) that focuses specifically on the coordination, development and investment to build campus capacity and capability now and for the future. The RISP complements and augments the current IT2020 plan. For the purposes of this strategic plan, research informatics is defined not only as all data collection, generation, computing, storage, analysis, visualization, access, and sharing in support of research but also the supporting infrastructure, technology, software, tools, services, policies, and governance.

The timing and the urgency of the RISP effort is motivated by the following needs and opportunities: (a) UCLA has been 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 (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 the development of a tangible action-oriented strategic plan, the planning process was divided into two phases: Phase One: Defining research informatics needs, barriers and obstacles that currently exist for faculty; Phase Two: Identifying a vision of themes that not only
set the context for first actions but also if pursued would enable UCLA to differentiate itself as a leader in research informatics in the next 5-10 years. Campus participation was extensive. Phase One was structured around five discipline-oriented committees involving over 70 faculty and staff. Phase Two involved individual discussions about future and vision with over 60 additional faculty. Using these initial results in a focused, take action approach reflects that RISP is neither all encompassing nor static, but rather a tool to initiate a campus-wide process that must be continued and iterative in order to be successful and transformative.

**UCLA’s Strategic Themes**

A primary objective of the planning process has been to establish the 5 – 10 year directional areas of emphasis or themes for the campus to focus near term decisions and to target investments of time and effort. 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 Sciences Research Group (HASIS), the Clinical Translation Science Institute (CTSI) Board, and additional faculty in the Social Sciences to identify the “big ideas” expected to influence research informatics in the next ten years.

Ten categorical themes emerged, capturing the UCLA pattern of an exceptionally broad, rich and large set of research activities at UCLA that had similar characteristics in multiple units. No one theme stands out as extraordinary when compared to other universities. However, the particular pattern of themes when contextualized with the UCLA’s faculty interests becomes UCLA’s roadmap for target investment. In other words, to be meaningful, the themes need to be viewed together as UCLA’s unique signature. It is only in this light that these ten themes and associated actions and investment become UCLA’s informatics strategic planning initiative.

While the ten themes are inextricably linked, Themes 1 and 2 primarily address the types of researchers that will be engaged in informatics, Themes 3 and 4 primarily address the type of tools, support services and infrastructure necessary to enable collaboration amongst researchers, and Themes 5 through 10 primarily address the various types and formats of data that will be collected, stored and analyzed in the growing field of informatics.

1. **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. The emergence of the citizen scholar will increase demand, access and contribution to shared data. 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.

2. **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.

3. **Research Ecosystem (rEcosystem).** An rEcosystem provides the institutional informatics connectivity and key shared resources that extend capacity; the tools and capability that form the institutional “nervous system”; and the rich functionality that brings distributed
resources, capabilities and expertise into a coherent capability, the links through which data can be moved and methodologies accessed; and institutionally leveraged infrastructure that builds individual capability beyond that which can be developed locally.

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

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

6. **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. 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.

7. **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. 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.

8. **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.

9. **Data Visualization.** Data visualization 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.

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, etc). 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. 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.

**Governance**

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

- 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 practice 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.
- Work with OVCR and OIT to perform an independent unbiased review and assessment of the oversight, organizational structure, scope, mission, and role of IDRE on campus, as well as provide recommendations on ways to optimize how that role interfaces and intersects with the wide spectrum of entities involved with informatics.

**Taking Action**

The following recommended actions 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.

1) **Transformative Recommendations**

A. Invest in fellowship-driven collaboratories in which advanced undergraduate, graduate, and doctoral students facilitate the application of computer science, statistical science, and computational informatics in areas that are either highly innovative or high yield in advancing cross disciplinary informatics research, experiences and tools with a diverse group of faculty.

B. Establish workshops for non-health sciences faculty and graduate student researchers to learn and discuss health data marts available for analytical use.

C. In partnership with the Institute of Digital Research and Education (IDRE) and the Clinical Translational Science Institute (CTSI), create a shared bioinformatics resource pilot 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.

D. 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 incentive 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.

Transformative Potential
The above recommended actions would: (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.

2) Catalytic Recommendations

A. Establish 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.

B. Institute a FacTech (Faculty Technology Awareness) 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.

C. Instantiate a Symposium Series for UCLA museums and campus departments on image archiving and the standardization and use of image metadata.

D. 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.

E. Create a speaker series of high profile external speakers to both engage and inform the campus of outside thinking and activities.

Catalytic Potential
The above recommended actions will: (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.

Next Steps
Review the strategic plan and recommended actions with the ITPB, Dean’s Council, Academic Senate, and then open to campus comment.