UCLA Institute for Digital Research & Education

# Empowering faculty-led innovation in digitally enabled research

2018-19 Strategic Planning Recommendation Report

# 2018-19 IDRE Strategic Planning Recommendation Report

*June 27, 2019*

Warren Mori, Jaime Marian, TV Singh, Lisa M. Snyder, Aaron J. Taber

*This report was endorsed by the IDRE Executive Committee on June 14, 2019.*

Anastassia Alexandrova, Chemistry & Biochemistry
Chris Anderson, Mathematics
Doug Bell, General Internal Medicine & Health Services Research
Andrea Bertozzi, Mathematics
Christine Borgman, Information Studies
Alex Bui, Radiological Sciences / Bioinformatics
Mark Cohen, Semel Institute / Resnick Neuropsychiatric Hospital
Eleazar Eskin, Biomathematics
Jacob Foster, Sociology
Alexander Hoffman, Microbiology, Immunology & Molecular Genetics
Lihua Jin, Mechanical & Aerospace Engineering
Chris Johanson, Classics
Jens Palsberg, Computer Science
Matteo Pellegrini, Molecular, Cell & Developmental Biology
Todd Presner, Germanic Languages & Digital Humanities
Pablo Saide, Atmospheric & Oceanic Sciences / Institute of the Environment and Sustainability
Marc Suchard, Human Genetics / Biomathematics
Joseph Teran, Mathematics
Marco Velli, Earth, Planetary & Space Sciences
Willeke Wendrich, Near Eastern Languages and Cultures / Cotsen Institute of Archaeology
Hong Zhou, Microbiology, Immunology & Molecular Genetics
Jasmine Zhou, Pathology & Laboratory Medicine

UCLA Institute for Digital Research & Education

*Learn more about IDRE at https://idre.ucla.edu*

# Contents



.

# Executive Summary

The Institute for Digital Research and Education (IDRE) was founded over ten years ago to bring together computer scientists, applied mathematicians, and statisticians with domain scientists and researchers from across campus. The intent was to create an environment for UCLA research groups to accelerate research discovery, find innovative solutions to complex problems, form interdisciplinary teams, and share experiences. IDRE has led or supported UCLA efforts in advanced computing, data analysis, and algorithm development in areas as diverse as bioinformatics, digital humanities, engineering, and computer simulations. As a measure and example of its impact, the IDRE Hoffman2 cluster – at 1,300+ nodes and 21,000 cores – is the largest campus compute resource within the UC system, servicing an increasing number of UCLA users, in excess of 1,500 in 2018.

In light of numerous recent changes at UCLA and in the national landscape, and considering the growing importance of digitally enabled research and cyberinfrastructure, IDRE has undergone a strategic planning process at the behest of its sponsors, Vice Chancellor for Research (VCR) Roger Wakimoto and Vice Provost, Information Technology (VP-IT) Jim Davis. Specifically, the process sought to:

- identify opportunities and emerging trends in digitally enabled research,
- understand the scope and common challenges of such research at UCLA,
- reassess IDRE's mission, structure, and role in support of such research,
- identify opportunities to leverage campus partnerships,
- expand IDRE's engagement with faculty across campus, including departments not traditionally associated with advanced research computing,
- assess how the growing demands of this research effects UCLA's educational mission, and
- assess how the campus can most effectively support interdisciplinary digitally enabled research at UCLA for the next five years.

The planning process included a series of meetings with individual faculty and key research groups, Departments, Institutes, Centers, and other campus entities. These interviews revealed that faculty across the entire campus are involved with digitally enabled research, and are internationally recognized in their respective disciplines. There is variation in the size and sophistication of research groups, along with levels of awareness about national resources, and overall experiences with advanced computing, storage, and networking. Peer institutions were also examined to learn about best practices and assess different institutional models for organizing and structuring digitally enabled research, including research support and services.

Based on the interviews, national initiatives, the rapidly changing digitally enabled research landscape, and the requirements of interdisciplinary research, this report lays out three recommendations for the campus to implement and for IDRE to lead in response to the UCLA faculty and researcher community:

1. Coalesce and focus campus efforts into three bold initiatives:
    a. Harnessing deep learning and next generation artificial intelligence,
    b. Harnessing the data and digitized information deluge, and
    c. Harnessing exascale, quantum, and cloud computing.

2. Spark research innovation by fostering interdisciplinary engagements and increasing IDRE's visibility.

3. Empower faculty through the advancement of campus capabilities and expertise for digitally enabled research, scholarship, and education.

The recommendations are detailed on the following pages. They describe three ambitious initiatives and supporting efforts that will enable UCLA to remain at the forefront of scholarly innovation, and engage and support faculty, researchers, and students. It is clear that the role of digitally enabled research and education at UCLA and across the country is growing and that the leading universities of tomorrow will be those who embrace this reality, enthusiastically address the challenges facing their faculty, and provide the supporting infrastructure that will spark innovation. Appendices to this report include details and narratives about the interviews and findings, peer institution research, and reference articles and resources.

## The Strategic Planning Process

The rationale for the formation of the Institute for Digital Research and Education (IDRE) is as strong today as when the Institute was first launched over a decade ago. Digitally enabled research continues to grow apace as data-driven exploration together with simulations take equal footing with theory and experiment as part of the scientific method. Digitally enabled research methods, including advanced computational and data science methods, are used and developed by UCLA faculty and researchers across the entire campus (North, South, and Medical). These methods are a critical component for solving many of today's research questions and grand challenges, and require knowledge in a "domain" area, advanced information technology systems (e.g., compute, storage, and networks), algorithms, methodologies, and software. Increasingly, addressing these research questions requires interdisciplinary teams that cut across traditional departmental boundaries and include domain researchers from across the entire campus, computational scientists, data scientists, information scientists, computer scientists, applied mathematicians, statisticians, librarians, archivists, and technologists. A critical challenge facing faculty is the complexity and sheer number of methods, software, algorithms, tools, and hardware used in digitally enabled research, making it difficult for a single faculty member or even a single research group to make proper decisions about cyberinfrastructure.

With these challenges comes the opportunity to spur innovation and facilitate breakthrough discoveries by coalescing and focusing the efforts of UCLA faculty and research groups together with support staff into strategic themes and initiatives in the areas of computational, data, and information sciences, and digital scholarship and creative activities. This also requires recurring campus investment and support. IDRE's broadly representative Executive Committee, faculty relationships, and close partnership with IDRE Research Technology Group (RTG) support staff and services is the natural campus entity to lead such efforts.

To underscore IDRE's current importance, during the past decade the IDRE Hoffman2 cluster has grown to more than 21,000 compute cores (.175 billion core-hours per year) with more than 1,500 individual UCLA users. In the past three years, IDRE has sponsored or co-sponsored 26 workshops and held 132 training sessions. In 2016, the IDRE Early Career Research Group was formed with four inaugural IDRE

postdoctoral scholars, all of who are now faculty members at other universities. The breadth of IDRE-related activities is also seen through 75 researcher profiles published in the IDRE newsletter.

At the national level, new initiatives reflect the complexity and importance of digitally enabled research, its intrinsic interdisciplinary character, the emergence of data science, and the recognition that academic needs for advanced hardware and cyberinfrastructure are continuing to grow. The Department of Energy, for example, is investing over $2 billion in an exascale computing initiative to accelerate scientific discovery through expansion of its computing ecosystem, and the Department of Defense is investing another $2 billion in what it calls third-generation artificial intelligence.[1] In addition, the National Science Foundation, other governmental agencies, and private foundations are increasing their portfolio of research in the use of advanced computing and data-driven discovery in many academic areas of research. Increased funding opportunities related to the development of computing resources, tools, and ecosystems are an acknowledgement that digitally enabled research requires advanced cyberinfrastructure including computers, storage, and networking; production-level software for simulation and data analysis (including artificial intelligence); skeleton software such as algorithms and libraries; new methodologies; access, efficiency, and delivery tools such as workflows and gateways; and new methods for analyzing large datasets.[2] Directly or indirectly, all UCLA faculty will eventually be affected by these funding decisions because of their impact on all aspects of advanced cyberinfrastructure.

## Background

IDRE's mission from near the time of its creation is to *"support, advance, and guide a campus-wide program to position UCLA as a world leader in research and education in computational thinking, using high performance computation, data visualization and data analysis of large data sets and databases."* A key goal is to foster a collaborative community of faculty, researchers, and technologists engaged in interdisciplinary digital research where successful completion requires expertise that cuts across traditional Department and School boundaries. This goal is reflected in IDRE's current vision statement:

> "*IDRE supports research and innovative scholarship that takes advantage of new technologies and encourages collaboration between faculty from different departments and disciplines at UCLA, the opening of new research questions, and the enrichment of the learning environment. The goal of this campus-wide collective is to make UCLA a world leader in high-performance computing and visualization research and education.*"

This report includes updated language for IDRE's mission and vision that reflects the current campus landscape and anticipates future trends.

IDRE currently relies on resources negotiated through a partnership with the Office of Information Technology (OIT) and the Office of the Vice Chancellor for Research (OVCR). This includes support from the IDRE Research Technology Group (RTG), a team of technologists within OIT that carries out the technical elements of the IDRE mission through working closely with faculty and researchers, and the development and maintenance of advanced computing and networking infrastructure, such as the

[1] See https://www.exascaleproject.org/ and https://www.darpa.mil/news-events/2018-09-07.

[2] NSF, for example, has framed its priorities around 10 Big Ideas, one of which is Harnessing the Data Revolution (other Big Ideas also significantly involve digitally enabled research). See https://www.nsf.gov/news/special_reports/big_ideas/.

Hoffman2 shared compute cluster. (See *Appendix A: Details about IDRE and Related Entities* for details about the purviews, mandates, and structures of those entities.) IDRE is currently organized into five programs that reflected the challenges and opportunities facing many IDRE-related faculty near the time of its creation: 1) Advanced Computation and Storage, 2) GPU/Many Core, 3) Pipeline to Leadership Class Facilities, 4) Statistical Computing/Data Informatics, and 5) GIS, Visualization and Modeling. This report also includes an updated organizational structure for IDRE, including new programs.

## Methodology

This report was commissioned by Vice Provost Jim Davis and Vice Chancellor Roger Wakimoto, and the strategic planning process was launched in the Fall of 2017. An internal planning group, led by current IDRE Executive Director Warren Mori, was convened to shepherd the process. The other members of the group were Jaime Marian (IDRE Associate Director), TV Singh (IDRE's Chief Computational Scientist), Lisa M. Snyder (OIT's Director of Campus Research Initiatives and Acting Director of the RTG), and Aaron J. Taber (OIT Project Manager). Additional input at several critical points in the development of this report was provided by members of the IDRE Executive Committee and members of the RTG. This included input on the scope of the planning process, participation in domain-specific interviews, and recommendations on peer-institution exemplars.

The internal planning group developed a list of questions to collect input from campus partners, focusing on the interviewee's perceptions of IDRE, use of IDRE resources and services, thoughts about possible enhancements to IDRE offerings and activities, and views on emerging research trends and roadblocks related to digitally enabled research. (See *Appendix B: Discussion Questions for Interviews*, and *Appendix C: Interviewed Campus Entities.*) The internal planning group interviewed a diverse mix of campus Institutes, Centers, and Departments that have experience with IDRE's services, or whose subject areas align with IDRE's mission, vision, and capabilities. An online form was provided to capture written input from faculty and associated members of these groups.

Over 45 peer institutions were examined to identify best practices and learn about different institutional models for organizing and structuring centralized research support and services. Following the methodology used in the 2017 "Research Data Support Recommendations" report,[3] schools examined are either in the top 25 of the 2016-2017 Times Higher Education World University Ranking, the top 25 of the 2015-2016 Times Higher Education World University Ranking, the top 25 of the U.S. News National Universities Rankings, or were suggested by the IDRE Executive Committee as having exemplary research and data management support infrastructure. Selected campuses were more exhaustively studied. (See *Appendix D: Peer Institution Research* and *Appendix E: Peer Institution Mission and Vision Statements.*)

The final element of the strategic planning research was an analysis of IDRE's strengths, weaknesses, opportunities, and threats (SWOT). This analysis was developed by the internal planning group and refined by the IDRE Executive Director. (See *Appendix F: SWOT Chart and Analysis*.)

[3] OVCR and OIT. *Research Informatics Strategic Planning (RISP) Initiative.* Report. 2013; OVCR, OIT, University Library, Chief Medical Officer for Clinical Research. *Research Data Support Recommendations*. Report. 2017.

## Findings

Regardless of interviewees' levels of engagement with IDRE activities and RTG resources, there was broad consistency with respect to higher-level recommendations related to campus support for digitally enabled research. Responses coalesced around the six key themes described below. These themes, in turn, informed the recommendations proposed in this report. (See *Appendix G: A Narrative on the Findings* for details about interviewee comments.)

### 1. Existence of a critical mass of faculty whose research is aligned with national initiatives

Within the Department of Energy, the National Institutes of Health, and the National Science Foundation, there are major initiatives including *Deep Learning and 3rd Generation Artificial Intelligence, Exascale and Advanced Computing,* and *Harnessing the Data Revolution*. In addition, the Office of the President recently announced an Artificial Intelligence Initiative. GPU (Graphical Processing Unit) computing is integral to both deep learning and exascale computing, and numerous faculty expressed interest in growing campus investment and support for GPU computing.[4]

### 2. Recognition of significant opportunities to leverage faculty expertise

UCLA includes an impressive array of faculty engaged in research that uses or develops advanced research computing methods, tools, and resources. Their diverse research topics fall within the terms of computational science (using advanced computing to solve complex problems), data science (extracting knowledge from data), information science (studying the storage and retrieval of information), and digital scholarship and creative activities (using digital technologies and tools to tackle complex inquiries into cultural and social problems). These faculty represent the entire campus. While many faculty are collaborating through IDRE in varying degrees, others are still not part of the IDRE community, which prevents them from benefitting from others and sharing their own expertise.

### 3. Exponential growth in campus-wide demand for access to computing resources

There was clear and overwhelming support for the Hoffman2 shared compute cluster, and a desire to see it grow in capacity. It remains widely used and demand is growing despite the rapid emergence of commercial providers offering cloud-computing solutions. Some faculty called for optimizing the queueing system, improving access to unused cycles, and exploring the ability to add nonconventional nodes and customization to the cluster, including GPUs.

### 4. Significant challenges related to data

Numerous groups across campus mentioned diverse issues related to the data lifecycle, data workflows, data security, data integration, and 'complex data' (though the latter was not uniformly defined). This was the case whether they use data generated from field measurements, sequencing, surveys, experiments, or simulations; or data gathered from private or public repositories. Challenges related to data include software, storage, curation, analysis and computation, and visualization. Groups mentioned a desire for sustainable and scalable storage solutions at the institutional level.

[4] As part of a campus proposal for an NSF Major Research Initiative for a GPU Computing Instrument, 81 faculty/researchers expressed interest in using such a resource and in developing new deep learning algorithms and applications.

### 5. Innovation and progress are hindered by the complexity of new technologies and methodologies

Faculty are aware of new technologies and cyberinfrastructure advances, but they do not have the expertise, time, or staffing to fully exploit the potential of these advances. Related to this challenge is a desire to have access to the latest computing hardware and peripherals. Examples cited include GPUs, FPGAs (Field Programmable Gate Arrays), 3D scanners and printers, and virtual reality headsets. Interviewees also expressed interest in new methods and algorithms (e.g., deep learning, scalable parallel solvers, agent-based methods, simulation, data science methods for extracting knowledge from data, topic modeling, network analysis), ways to develop and maintain software (e.g., software optimization and engineering, version control, issues porting software to new hardware), and emerging research tools and applications (e.g., Jupyter notebooks, cloud computing and storage, science gateways like Galaxy, augmented reality, Tableau, RSpace, Dedoose). Finally, while some researchers on campus are engaged with these technologies and methodologies, it was noted that there is no campus-wide mechanism for sharing use cases and best practices.

### 6. The value and impact of IDRE and the Research Technology Group

Most faculty and researchers interviewed are aware of IDRE's Advanced Computing Program and the Hoffman2 cluster, and many associate IDRE with the services provided by the IDRE RTG. There was broad appreciation of the RTG's unique expertise, and faculty expressed interest in increasing the number of RTG technologists, and better clarifying the process through which faculty engage them in research support. Researchers across campus benefit from IDRE's training classes and workshops, would like to see them expanded, and there was support for IDRE to focus on early career researchers as an engine for collaboration.

## New Vision, Mission, and Organizational Structure

This planning process made clear that IDRE's vision, mission, and organizational structure need to be updated to better reflect the new landscape of digitally enabled research. In order to leverage the unparalleled range of faculty knowledge and active research in digitally enabled scholarship, **IDRE's vision should be "to sustain UCLA as a world leader in the innovative use, application, and development of computational, data, and information sciences, and digital technologies that drive cultural and societal change."** Achieving this vision will require that faculty work together and form interdisciplinary teams to solve complex problems, spark innovation, and accelerate research discovery and knowledge production at UCLA. Similarly, IDRE must advance UCLA's educational mission by facilitating innovations in digitally enabled pedagogy and by developing curricula that responds to the rapidly evolving research landscape. To achieve this goal, IDRE must become a top priority of the campus administration, and an investment must be made in advanced cyberinfrastructure and creative research support solutions and services.

Since its inception, new domain specific Institutes or Centers have been created at UCLA. As a result, the mission has shifted from supporting specific domain areas to the use and development of advanced cyberinfrastructure. **IDRE's new mission is "to lead a campus-wide effort on innovative research and education on the use and development of advanced cyberinfrastructure methods, tools, and services. This involves bringing together researchers working in domain, computational, data, and information sciences, and digital scholarship and creative activities."**

IDRE will be organized into four interlocking pillars focused on Computational Science, Data Science, Information Science, and Digital Scholarship and Creative Activities with crosscutting activities common to each program (see figure below). As seen in the exemplars, some universities have institutes with a similar organizational structure, others only have an institute focused on either computational or data science, while still others only have technology services centers or domain-specific centers. Based on UCLA's breadth, an institute like IDRE makes more sense as there is significant overlap in faculty interests between computational, data, and information sciences, and digital humanities, digital social sciences, and digital media and arts. In fact, data and information sciences function as a critical link between south and north campus in the figure below. Additionally, the creation of the Digital Scholarship and Creative Activities program recognizes the importance of engaging all faculty, including those who may not necessarily identify their research (including creative activities) as strictly aligned with Computational, Data, or Information Science. This program includes and advances the work of the IDRE-HASIS (Humanities, Arts, Social, and Information Sciences) group that focuses on integrating research, education, and outreach endeavors specific to those domains.

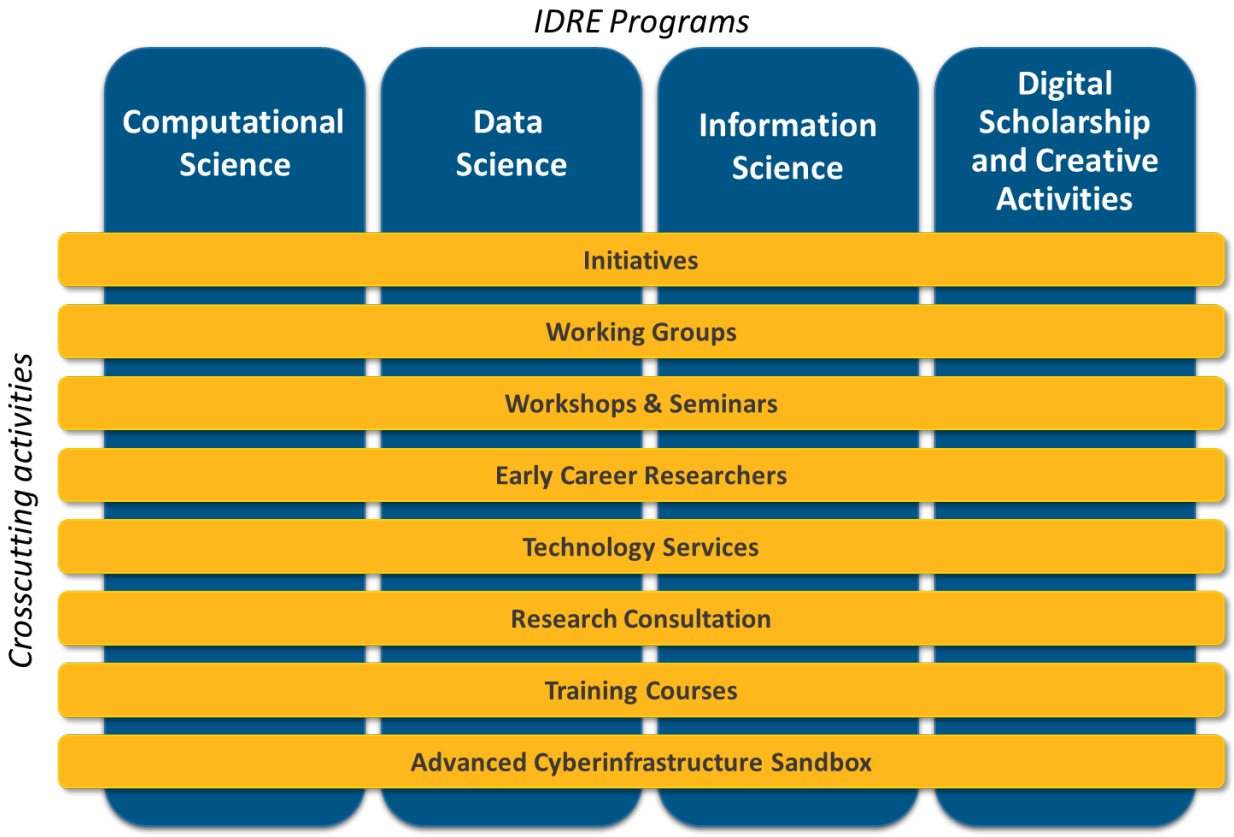


*Figure 1: IDRE's envisioned programs and crosscutting activities*

# Recommendations

In order for UCLA to fulfill its potential as a world leader in digital research and education, it must create an environment for faculty and scholars to focus their collective efforts on ambitious and timely initiatives, share experiences and lessons learned, and build successful collaborations. This report describes three key recommendations for IDRE to implement to create such an environment. These recommendations also lay out a process for IDRE to continually examine and realign its focus to changing campus and faculty needs, new technologies, shifts in the national funding landscape, strategic areas for new faculty hires, and areas for enriching and coordinating education in digitally enabled scholarship. Taken together, they provide a roadmap for IDRE to achieve its vision and serve its mission.

## Recommendation 1: Coalesce and focus campus efforts into three bold initiatives

The campus should focus on – and IDRE should lead – a small number of critical campus initiatives that will capture the imagination of the entire campus, ignite faculty and researcher innovation, leverage common efforts on campus, and broadly inform the campus of the impact of emerging new digitally enabled research topics. During the planning process, a few broad areas of common interests emerged that have shaped the following three initiatives.

### 1. Harnessing deep learning and next generation artificial intelligence

*Goal: Capitalize on the growing interest in AI (artificial intelligence) and DL (deep learning), raise the research productivity across campus, find grand challenge research questions for new collaborations, develop next generation AI and DL algorithms, develop homegrown software, disseminate breakthroughs in next generation AI algorithms and methods, and critically examine the social, cultural, and ethical implications.*

The national news and academic journals are replete with examples of how AI and DL are impacting everyday life, society, and cutting-edge research. Although many of the core ideas behind these methods have been around for more than a decade, increases in computing power and new ideas for applying these methods are transforming research at UCLA. There is no current coordination of this type of activity on campus.

### 2. Harnessing the deluge of data and digital information

*Goal: Create a culture of innovation, form grand challenge research themes, develop a transparent set of policies around the responsible use of data, data privacy, storage, and cybersecurity, and build a coordinated curriculum for preparing a 21st century workforce for a world brimming with data and digital information.*

Data is everywhere and faculty from across campus face challenges with acquiring, using, and preserving it. It is important to emphasize that challenges with data (including software) at a university fall into two separate yet equal areas. The first is how to enable fundamental research discovery through the use of data and the innovative development of new digital tools and technologies that advance the creation and analysis of varied data types across campus. The second is how to enable ease of access, aggregation, storage, curation, retrieval, computation, conversion to information/narratives, and the preservation of data/information (including privacy

and regulatory issues as well as broader ethical issues). These issues are at the core of data science, information science, and digital scholarship.

### 3. Harnessing exascale, quantum, and cloud computing

*Goal: Increase access to national user facilities, including exascale platforms, provide local GPU/accelerator hardware resources, foster the development of software that can effectively utilize GPU/accelerators, communicate to the campus future trends in next generation advanced computing architecture, coalesce efforts in quantum computing, and ensure access for UCLA researchers to an ecosystem of national, local, and cloud resources.*

The United States is committed to enabling exascale computing within the next five years.[5] Such a computer is a very complex, billion dollar piece of equipment, that will be used for simulations in physics, chemistry, biology, engineering, atmospheric sciences, and for AI and DL in disciplines including predictive medicine and health sciences. Writing software that can use such computing is a daunting task. Currently, the path towards exascale is based on using GPUs (100,000 of them), which is also the hardware driving deep learning. UCLA researchers are therefore requesting and requiring access to GPUs and other similar advanced computing capability, whether on premise or via commercial cloud computing services.

## Recommendation 2: Spark research innovation by fostering interdisciplinary engagements and increasing IDRE's visibility

IDRE should foster innovation by facilitating interdisciplinary collaborations between faculty, researchers, and other campus units/centers/institutes from across the entire campus and by providing the mechanisms to develop, discuss, and learn new methods, best practices, and trends in their use of advanced research cyberinfrastructure. This important goal for IDRE is to ensure that faculty and their research groups interact, share knowledge and best practices, collaborate on research, and develop strategic priorities for campus resources. This will be achieved through crosscutting activities, including those described below. As extramural funding opportunities continue to highlight multidisciplinary teams, IDRE must provide ways for faculty to create successful relationships among colleagues. Increased interactions with faculty and researchers – including those early in their careers – will help IDRE continue to shape and refine its initiatives, programs, and activities.

### A. Faculty working groups

Faculty-led working groups should be formed to focus on timely topics and common interests on digitally enabled research and scholarship. Activities should include birds-of-a-feather lunches, forums, and working groups focused on IDRE's initiatives. The topics of the working groups should be broad enough that there is interest from a critical mass of research groups, yet narrow enough that the groups' efforts can be focused. Working groups would naturally identify themes for workshops, invited speakers, visitors, and tools for experimentation within the advanced cyberinfrastructure sandbox (described in Recommendation 3), as well as providing a space for collaboration to seek extramural funding. The working groups can also provide a mechanism for

[5] Exascale computers will perform $10^{18}$ (a billion billion) floating point operations per second and have $10^{17}$ Bytes of memory (100,000 TBytes).

faculty to interact with RTG technologists, and provide input for activities run within a program or in support of IDRE's initiatives.

## B. Faculty-led workshops and seminars

IDRE should develop and leverage its community of interdisciplinary researchers and use its workshops, seminars, and social events to review its strategic plan continuously. Over the years, IDRE has offered a limited number of faculty-led workshops (similar in style to a short IPAM program) as well as more training-style workshops led by outside advanced computing centers or experts. These have been well-attended and impactful, and should be offered on a regular basis. IDRE has also hosted a number of invited speakers. These activities should be leveraged by soliciting partners both on and off campus to co-host events, including invited speakers and follow-on sessions to larger events (e.g., a UCLA-centric workshop following an IPAM conference). IDRE should sponsor a minimum of two workshops per quarter and host an equal number of invited lectures.

## C. Early Career Researchers (ECR) and postdoctoral scholars

The IDRE *Early Career Researchers* group should be expanded and more widely publicized, including the solicitation of incoming graduate students and postdoctoral scholars each Fall. The ECR brings together graduate students, postdoctoral scholars, and other researchers early in their post-graduate careers from a variety of disciplines and methodological perspectives interested in advancing cross-disciplinary research in computational and data sciences at UCLA. Regular activities like meetings, internal and external seminars, social events, and poster sessions should occur regularly. Expanding the virtual ECR community through social media and the sharing of relevant departmental events should also be explored. The ECR's Postdoctoral Fellowship Program should leverage existing postdoctoral programs at many of UCLA's domain-specific Centers and Institutes, and explore efforts to create a shared postdoctoral program that would provide more visibility to the postdocs, as well as organically facilitate collaboration between IDRE and domain-specific Centers or Institutes.

## D. External partnerships, visiting scholars, and connections to industry and society

IDRE should proactively develop partnerships with other advanced computing centers that serve academic research in computational and data science, including those at peer universities and at national laboratories. In addition, IDRE should explore opportunities to work with partners in industry, community organizations, cultural institutes, and the non-profit sector. The goal of these partnerships is to increase expertise and awareness in emerging technologies, to share best practices and expertise, and to develop collaborative projects with public impact. They would also improve the visibility of IDRE and the IDRE RTG, and possibly lead to partnerships for new funding opportunities.

To reinforce ties with external partners and peer institutions, IDRE should explore the development of a visitors/sabbatical scholars program aimed at interdisciplinary research. Even at a world-class research university such as UCLA, there can be shortfalls in local expertise for fully developing interdisciplinary research teams, leading to non-optimal extramural proposals. While the campus' long-term solution may be the strategic hiring of new faculty, a

visitors/sabbatical scholars program is a quick and cost effective way to invest in areas that could have a large impact to UCLA.

### E. Promotion and outreach

IDRE must expand its promotion efforts in order to inform faculty and administrators of the opportunities and impact it provides, and to increase the number of faculty from across the entire campus who participate in IDRE activities. In order to maintain the support of the faculty, Deans, and its sponsoring entities, it is important that IDRE champion its activities and successes. The development of an annual report, marketing collateral, and continual outreach are essential parts of this effort. Additionally, IDRE should provide template language that researchers can use when writing interdisciplinary proposals that involve the use of advanced cyberinfrastructure. IDRE should also seek to be cited as appropriate in publications.

## Recommendation 3: Empower faculty through the advancement of campus capabilities and expertise in digitally enabled research, scholarship, and education

To better support the growing complexity of faculty research, IDRE must advocate for the expansion of research support services, professional staff with technical expertise, advanced research computing resources, and enhanced cyberinfrastructure. The Institute must also help shape the cohesive institutional response to common research challenges that cannot be met by any one Department, School, or Division alone. IDRE can help faculty break through obstacles that they increasingly face when carrying out research that develops or requires the use of advanced cyberinfrastructure, demands interdisciplinary collaboration, or involves new and experimental technologies. Moreover, IDRE is also well positioned to explore novel and pioneering technologies, and subsequently make technical recommendations for individual faculty, research groups, and campus administrators. With institutional support, IDRE can advance the University's research mission, and make UCLA a national leader in the use and development of innovative digitally enabled scholarship.

### A. Campus advanced computing capability and research support expertise

IDRE must advocate for the continued expansion of and support for UCLA's research computing environment, and develop a five-year roadmap for meeting the campus' growing research support needs. The local Hoffman2 cluster, along with national facilities and IDRE's initiative to explore cloud and exascale options form a natural broad ecosystem of computing and storage options for faculty. (See *Appendix A: Details about IDRE and Related Entities* for technical details about the current configuration of Hoffman2.) The Hoffman2 Users Group (HUG) – made up of faculty, other Hoffman2 users, and RTG technologists – should reconvene and meet regularly to discuss and recommend potential hardware/service upgrades to UCLA's research cyberinfrastructure.

IDRE should create and foster a more coordinated research support experience by building on existing RTG expertise and leveraging partnerships with other campus entities and peer entities at national labs and universities. In the interviews, faculty from across campus repeatedly pressed for access to staff expertise that supports their research. The current stable of RTG technologists are valued and should be enhanced through strategic new hires, as well as through the

temporary or permanent co-location of experts from other related Institutes, Centers, and School/Department service groups on campus. IDRE should also explore creating a 'research facilitators' program to guide faculty to campus, third party, and free national resources, and extramural funding opportunities that best suit their needs. This will also help IDRE faculty identify future funding opportunities and collaborations, and provide a coordinated user experience with a high degree of technical and customer support. This concept is growing at peer institutions, led, in part by the Advanced Cyberinfrastructure – Research and Education Facilitators consortium (ACI-REF).[6]

## B. Strategic partnerships and growth of the Research Technology Group (RTG)

In order to solidify IDRE's leadership role across the campus in driving research and educational innovation, strategic partnerships with other units, such as the Library and local Institutes/Centers, are necessary. This can be achieved by growing the footprint of the RTG in both existing areas of strength (i.e., HPC, GIS, visualization, modelling, and statistical computing) and developing new areas of expertise attuned to needs across north and south campus. Continued investments in first-rate staff will be critical for IDRE's next decade of success.

## C. Advanced cyberinfrastructure experimental sandbox

IDRE needs to create an experimental environment where faculty and research groups can explore solutions to the challenges they face with the rapid emergence of new cyberinfrastructure tools, algorithms, software, and hardware. Such an 'Advanced Cyberinfrastructure Experimental Sandbox' would benefit faculty from across campus and facilitate interactions among members of their group, other researchers, and RTG technologists. In addition to streamlining research access to new tools, algorithms, and hardware, this sandbox would help researchers determine when best to invest time and resources into new cyberinfrastructure, and provide a template for others on campus facing similar challenges. The topics to be investigated will be informed by IDRE's initiatives, working groups, solicited suggestions, and input from the RTG.

## D. Trainings courses

The roster of IDRE and RTG training courses should be expanded in response to changing campus needs and interests. This includes elevating introductory-level training courses, and developing courses that address more advanced and complex research challenges. An advisory committee of IDRE faculty, early career researchers, and RTG technologists representing various campus constituents should recommend a cohesive list of training courses to be offered each academic year, and make a concerted effort to compliment, coordinate, and leverage the offerings from campus partners. These training opportunities can be promoted to the campus through the recently launched Workshops@UCLA portal (http://workshops.ucla.edu) which aggregates course listings from IDRE and partner organizations across the campus.

[6] See https://aciref.org