

Research Data Working Group Roadmap Update

Last modified August 7, 2023

Campus Feedback on the RDWG Roadmap

Following the Winter Quarter 2023 release of the Research Data Working Group's roadmap for expanding research data services and infrastructure at UCLA, the RDWG steering committee launched a campaign to gather feedback on the roadmap from academic leadership campuswide, appropriate Academic Senate committees, the Institute for Digital Research and Education faculty board, DataX representatives, faculty, and IT personnel. Overall, representatives from the RDWG have, to date, conducted 33 meetings with UCLA leadership including deans, associate deans for research or the equivalent, chair's and IT leadership from nearly all schools across campus.¹ The RDWG also conducted four public faculty forums and held meetings with relevant committees from the Academic Senate. This report summarizes those conversations and outlines next steps for moving the effort forward.²

Overall, the participants in these feedback-gathering sessions were strongly supportive and appreciative of the work involved, which they rightly interpreted as being on their behalf. The conversations reinforced the challenges articulated in the roadmap and the extent to which the campus is behind peer institutions in terms of research data services, support, and human and technical infrastructure. The feedback shared common themes:

- That the campus needs to do something to better support research.
- That we need to work together to address and fund our common challenges.
- That we need cost-effective, efficient, and simple solutions that minimize redundancies and duplicative effort.
- And that we need to streamline the bureaucratic processes associated with research activities.

The following discussion focuses on those trends and requests that have the greatest relevance to research data and the content of the roadmap.³

¹ See Appendix 1 for a list of the meetings and presentations.

² Notes taken during the meetings and presentations were analyzed by four reviewers, two that participated in each session (RDWG co-chairs Todd Grappone and Lisa M. Snyder) and two that were basing their interpretations solely on the written notes (OARC's Eleanor Koehl and Pablo Morales). The reviewers independently produced lists of themes and needs before meeting to discuss and distill their thoughts for this document. For sessions where there were notes from multiple notetakers (e.g., from the faculty forums), the different versions were packaged together for review.

³ While the conversations were structured to garner feedback on the RDWG roadmap specifically, participants often expressed frustration with aspects of how UCLA supports the research lifecycle generally. Purchasing and the third-party risk management process (TPRM), for example, were recurring targets though both are arguably outside the scope of the RDWG.

Concerning Trends

Specific situations that are particularly worrisome from a campus perspective emerged from the conversations:

1. **Faculty are seeking out research support options at other institutions because of challenges and limitations at UCLA.**

This trend manifested in multiple ways. Faculty are actively considering pushing grant administration to collaborators' institutions. (One faculty executive committee chair reflected on their own research saying, *"We're thinking about the follow up to this grant, and we're saying let's do the next one at Notre Dame."*) Researchers are also using partners' institutional compute and collaboration resources and/or relying on services developed at other institutions (e.g., the UC Davis Health Cloud Innovation Center, Wharton Research Data Services, San Diego Supercomputing Center, the Online Cultural and Historical Research Environment (OCHRE) at the University of Chicago).

2. **Incoming and potential faculty perceive UCLA's research infrastructure as being behind that of peer institutions.**

Challenges with recruitment, retention, and onboarding were routinely mentioned as ongoing issues for the academic leadership teams. The lack of support resources was cited by one former dean as the reason a star recruitment failed. Similarly, successful hires find the research environment lacking. (*"[We've got] faculty coming from other institutions and they say, 'Hey, I'm coming from X and it was so easy there.' It feels like they've slid backwards in terms of their storage needs."*) There was acknowledgment that campus investment in research is long overdue. This perception is validated by the peer institution research conducted by the RDWG during the roadmap development process. This gap between UCLA and peer institutions will continue to grow in the absence of investments in research data services and infrastructure.

3. **Researchers are struggling with limitations and demands.**

The limitations of UCLA's current research environment means that faculty are facing challenges securing staff expertise, working with large data sets, manipulating and transferring data, accessing compute power and storage, and meeting data-related funder requirements. (*"We're putting out fires every day. There's a lot of money being lost as well. Delays. NIH gets upset. Other institutions pull back funding. [It] becomes very real both for PI and the School."*) One PI described shifting staff on a grant from research to data management because of the funder requirement for publicly available datasets. Another unit is struggling to provide adequate grant support staff. (*"One challenge is getting the right amount of staff people to handle post awards. [It's] hard to find them. They're jumping around a lot, getting picked off by different departments even on the same campus. It's hard, especially for multi-institutional grants. It's a people problem."*)

4. **Members of the campus community are subverting the structures/systems in place at UCLA because they are perceived as being too onerous or lacking.**
Examples include buying software on their own rather than following the TPRM process, buying and hosting equipment on a researcher-by-researcher basis and exposing the campus to security vulnerabilities. (*"The whole purchasing and invoicing [situation] is a big issue. A lot of the large vendors aren't filling our orders because we're not paying our bills on time."*)
5. **Duplicative effort is wasting time and money at both the faculty and academic leadership levels.**
As one dean noted, *"I am always stunned at the 17 places where the same things are happening and encourage you to make sure that those overlaps are foregrounded and addressed."* There was general acknowledgement that faculty are spending considerable time assembling basic resources with effort spent buying equipment, figuring out data commons-like environments (e.g., data, shape files, shared data needs), building closeted computing systems, and fielding requests for and/or buying large storage drives. (*"If we don't have the infrastructure in place, that will put a lot of obstacles and barriers to allowing [our] work to take place. It's preventing people from engaging with the research. Spending time and energy making this stuff work. Parallel to when we get grants and don't have the infrastructure, people spend a lot of time trying to figure out how to get things done."*) *"People have to reinvent the wheel because there isn't infrastructure in place."* *"When we recruit in data sciences, it's costing us a lot of money in start-up costs. It's breaking the bank."*)
6. **There is significant faculty frustration with the way UCLA currently operates.**
This trend was variously attributed to a 'vision problem' and bureaucratic challenges. On the latter point, the argument is that UCLA's bureaucracy is NOT moving at the pace of research and is, indeed, getting in the way of innovation. There is the sense among the faculty that the campus is working against them, that there's bureaucracy in place, but it's not helping people do what they need to do. (*"It feels like there are policies in place, but there's an inflexibility to take something case-by-case and it requires eight people to find a solution. It should be easier than that."*) *"NIH awards are made to UCLA and not UC Health, but UC Health requires nonsensical oversight that, more often than not, is not applicable to non-human data collection and or non-patient behavioral science data."* *"Any time we can make it less frustrating, smoother – any time the faculty do well, the whole university does well."*)

Meeting Campus Needs

The conversations with the faculty and academic leadership teams reinforced the need for the main roadmap elements and the importance of working together as a campus to develop simple solutions that work. (*"We're pretty decentralized and everyone is off doing their own thing. If you can get us thinking about one campus...make it more cohesive so everyone understands we're*

in a shared environment.”) Concerns were also expressed, primarily about cost, centralization generally, and the impact on already overburdened local IT teams. Certain items from the RDWG resonated more strongly than others and the table below identifies those that were mentioned repeatedly in the discussions and will be prioritized during the upcoming implementation phase of the effort.

Recurring Requests	Select Quotes	Roadmap Solution
Cohesive information about resources on campus	<i>“Most [challenging] is trying to find resources on campus as well. Provide a document on resources on campus. UCSF updated an annual list of resources that one could use to cut and paste into proposals. Having one source for campus with that information would be really great.”</i> <i>“Explaining to everybody what resources are out there and how to access them is critical.”</i>	RDR mission to create resource catalog; should involve existing research concierge efforts
Computing, GPUs, and cloud options	<i>“Everybody sees the need for additional computing power with modern equipment.”</i> <i>“We need Tensor Processing Units and support for that.”</i> <i>“[We need] high-performance computing, the right environment for [our] students to have the tools they need, and fast storage directly attached.”</i> <i>“Say you’re working on a massive database. You need lots of machines and need lots of data. [And it will] run for a long time to train a new model. That’s foundational CS research. For that you wouldn’t need something local; you might try something locally and push to cloud later.”</i>	Expansion of existing research computing infrastructure
Storage	<i>“Storage is a huge issue. Chairs would email me about storage – backup storage, collaboration space. ... In the division they offer a generous baseline; establishing that baseline is low-hanging fruit. Everyone needs storage space: Box, Google Drive, Synology, research data backup, co-storage, allocations [with] permissions.”</i> <i>“They’re putting [data] all over the place. The lack of a clear set of choices to go to [means] that everyone is finding their own way, so keeping this secure is impossible.”</i>	Bruin Research Data Hub (NOTE: Storage is a particularly critical issue at this time considering the ongoing Google Drive containment effort)

Recurring Requests	Select Quotes	Roadmap Solution
Specifically, secure data storage (P3/P4)	<i>“At this stage, they don’t care where [their data] goes, They need to go to a place that’s OK, they can describe that in the proposal, and they can easily get it there. Convenience and simplicity.”</i> <i>“[One of my researchers] needed a FISMA solution and had to go to Irvine or Riverside for that. SDSC may have been the solution. We don’t have a place that meets all levels of security.”</i> <i>“They’re hungry for a central place that will check all the boxes, that’s quick and easy. Something that is easy.”</i> From an IT director: <i>“We can’t currently provide HIPAA-compliant storage. We’re hoping for help for folks with a foot in the med center that work with human subjects. If we can’t help them, they’ll roll their own and it’s not secure and not scalable.”</i>	Possible in either the Bruin Research Data Hub or through the expansion of existing research infrastructure
Solutions to streamline collaboration, whether on campus or with external partners	<i>“I get from our research centers requests for a platform ... for a way to circulate documents and collaborate.”</i> From the lead PI on a large NIH grant, though not the corresponding investigator: <i>“[As a result,] I can’t access and use the PI Portal. I can’t even use Box on South Campus. So when we talk about inter-institutional collaboration, it’s as if we’re two different institutions.”</i>	Part of RDR mission and baked into the Bruin Research Data Hub
Safe ways to access and share files and data	<i>“Research groups come to us and they’re looking for storage locations and collaboration, specifically collaborations [involving] data.”</i> <i>“[We] need ways to share large files. [Something] safe yet accessible to people outside UCLA. [Our files are] too large to share and people are not able to access our Box [folders].”</i>	Bruin Research Data Hub
Cohesive strategy for campus software licensing	Following a complaint about purchasing: <i>“The short-term solution is that people buy the software with their own money or get free versions online, which is a huge security risk. The policy as it is being implemented is a detriment to researchers and the campus. It makes things worse rather than better. Bad for security and researchers.”</i>	Part of RDR mission and baked into the Bruin Research Data Hub
Expertise related to all aspects of the data lifecycle	<i>“Our investigators will need more statistical management of these datasets. That’s going to be frustrating for investigators down the line.”</i> <i>“I’m thinking about data security. When this [roadmap] is all laid out, [we’ll need] a dedicated security team and dedicated admin team that understands the data that’s</i>	Expansion of existing research data services

Recurring Requests	Select Quotes	Roadmap Solution
	<i>flowing in and out, to work with this new service to support it. That would make researchers believe their data are safe here.”</i>	
Data services (e.g., metadata, publishing, reproducibility)	<i>“NIH must know that we don’t have the capacity to do this. We’ve been in conversations with our program officers. They ask us what campus resources we have, [and] we tell them we’re responsible to do it for the grant. ... Our program officers are surprised that we don’t have any support in this. That we have to provide the capacity as a team.”</i>	Expansion of existing research data services

Expanding the Roadmap

One outcome of the extensive review process is that it confirmed that the original roadmap document accurately represents the research needs of the campus. Additions were, however, identified in the course of the interviews though none that require modification of the discussion draft currently posted on the ORCA website. To simplify analysis and prioritization, a spreadsheet was created with rows for each of the tasks called out in the RDWG roadmap.⁴ The task numbers reflect the numbering in the roadmap document. Once the roadmap elements were in place, new items were added based on recommendations from the faculty forums and one-on-one interviews with the academic leadership teams. Rather than task numbers, these entries are identified in the spreadsheet by the source of the request.

Two items have been added that the RDWG organizers felt were implicit in the original document but not explicitly numbered:

1. Develop a security and compliance program integrated with research, and
2. Streamline interactions between the campus and the health sciences.

Seven items were added in response to campus feedback. These new items complement the original roadmap tasks:

1. Develop a program focused on proactive outreach to faculty regarding available resources, best practices, and compliance;
2. Virtual machines (as part of the Bruin Research Data Hub);
3. Containers (as part of the Bruin Research Data Hub);
4. Identify and implement co-location options;
5. Identify and implement hosting options;

⁴ See Appendix 2 for the RDWG Implementation Task Analysis and an explanation of the spreadsheet data. This task list includes an initial prioritization based on the campus feedback described herein.

6. Research business analytics; and
7. Faculty profiles.

Finally, five items were added to the implementation task list that are known to be part of the strategic planning effort or ongoing within IT Services:

1. Team to assist with large-scale research proposals,
2. Improve proposal tracking and other grant reporting requirements,
3. Add additional search capabilities in PATS,
4. Network unification, and
5. Modern Identity and Access Management (IAM) system for UCLA.

Moving Forward

It is clear from our conversations with faculty and the academic leadership teams that UCLA needs to take action to better support research and to do so in a cohesive, robust, and thoughtful manner.

While not an explicit goal of the RDWG, success in implementing the research data services and infrastructure recommended in the roadmap will also improve the likelihood of success for other campus initiatives. The vision of DataX, for example, hinges on the availability of a cohesive research support network complete with compute resources, staff expertise, and data collections and services. In its description of the DataX Support Network, the *Report of the DataX Strategy Task Force: A Template for Transportation* calls for staff at both the local and central levels to “function as a kind of “circulatory system” for knowledge and techniques relevant to DataX [and] enable broad involvement of researchers and creators across campus.”⁵ The report further describes staff capable of supporting the data-related activities (e.g., data acquisition, curation, manipulation, access, and management) and points to key central units (e.g., the Library’s Data Science Center, the Research Technology Groups in OARC) that will work with discipline and departmental support staff to “facilitate faculty and student use of existing data science techniques, and offer workshops, tutorials, mini-courses, or bootcamps where appropriate.”⁶ The RDWG roadmap includes the hiring of additional staff with data and computing expertise that will be able to function in these capacities. Moreover, the roadmap advocates for the development of the research infrastructure (e.g., compute and storage capabilities, collaboration tools, data-related software, networking) necessary to support the campuswide collaborations that will ensure the success of the DataX initiative.

Moving forward with implementation will require prioritization, continuing to build campuswide consensus on the initial list of chosen solutions, and building trust that the campus can successfully address our current research data services and infrastructure challenges. It will be

⁵ Page 46 of the DataX report downloadable from <https://ucla.app.box.com/v/UCLA-DataX-2022-Report>.

⁶ *ibid.*

critical to retain current solutions that work and fix those that don't, and remember that our research support network moving forward has to prioritize staff support as well as technology. Our next generation solution set has to foreground ease of use and simplicity, accommodate the needs of different disciplines, incentivize best practices, and be built with faculty input to ensure that their needs are being met. A mix of solutions will always be required to address different disciplinary needs; our goal is a suite of central resources that will unite all disciplines with common services and common understanding of available research tools.

Leveraging research- and data-related resources and services available nationally and at the UC level has always been a part of the roadmap vision. UCLA researchers are currently accessing NIH datasets housed on commercial cloud platforms as part of the STRIDES initiative, receiving compute allocations on national class facilities through NSF's ACCESS program, depositing their data and publication datasets with domain-specific repositories, publishing open-access copies of their work in the UC's e-Scholarship Repository, and securing services from third parties or at other institutions (e.g., individual agreements for data backup with CrashPlan, SDSC's 'Sherlock' data storage service). At the time of this writing there are also two ongoing UC-wide efforts that may provide solutions to complement the RDWG roadmap, the first being a re-launch of the Research Data Backup RFP effort to identify third-party research data backup solutions and the second being the possible system-wide expansion of the UC Santa Cruz Open Source Program Office (OSPO) pending funding from the Sloan Foundation. The RDWG roadmap recognizes the importance of aligning local services with these and other critical external-to-UCLA resources and expanding their use by the UCLA community. The availability of these resources does not, however, negate the need to develop the local services advocated for in the roadmap. The feedback sessions made clear that faculty needs span the research lifecycle from raw working data to publication datasets in need of curation and long-term preservation, and the best institutional response is to provide them access to a quality research environment. This is especially true when the available repositories and external resources are inconsistent across disciplines, do not support data needs across the full research lifecycle, and do not address many of the most challenging issues (e.g., collaboration, secure data storage, local data services) that are critical to the research function.

In respect to data publication, predicting the future of data repositories is difficult. One might imagine endorsements from the professional associations would establish a finite number of dominant repositories or that our major funding agencies would create national repositories similar to those in other countries (e.g., the Australian Data Archive), but does not seem likely in the near-term. In a presentation at the Spring 2023 Membership Meeting of the Coalition for Networked Information (CNI), Martin Halbert of the NSF very specifically addressed this point: "I can attest that not all agencies are going to stand up their own data repositories. NSF is not." "We ... encourage data to be repositied in any of the many long-established disciplinary [and] institutional repositories that NSF has often funded over the years and made major investments in. We do not want to undercut any of those investments."⁷ Among the responsibilities of the

⁷ Coalition for Networked Information (CNI) Spring 2023 Membership Meeting, held April 3-4, 2023, in Denver CO. See <https://vimeo.com/818398169> (~at the 40-minute mark)

proposed research data office is directing researchers to the most efficient, cost-effective, and relevant home for their research data. The goal is to ensure that UCLA research data are widely and publicly available or safe and secure as appropriate, and to provide support to researchers as they navigate the constantly evolving landscape of research data repositories and requirements.⁸

Unless otherwise directed, the RDWG effort is going to move forward in the following fashion:

1. Formally dissolve existing RDWG entities

Moving into an implementation phase will require targeted expertise and suggests the dissolution of the existing executive sponsor group, the steering committee, and the four subcommittees.⁹ The RDWG co-chairs will send thank-you letters to all current participants explaining how the effort will move into the implementation phase and letting them know that they may again be called into service depending on the final prioritization of the roadmap elements.

2. Secure a new charge letter from VCRCA Roger Wakimoto and EVCP Darnell Hunt

This new charge is to be based on the prioritization of roadmap elements that resulted from the RDWG meetings with faculty and academic leadership teams and overlaps with the UCLA Strategic Planning effort. It should include formal expectations for the convening of cross-cutting campuswide implementation teams as appropriate to address each prioritized roadmap element, the submission of funding requests, and timeframes for completion. Lisa Snyder, Todd Grappone, and Mary Okino will function as implementation coordinators and will be joined by a faculty member that understands UCLA's research environment. This team will coordinate the implementation teams and work to make progress on roadmap elements that are deemed possible to move forward immediately, whether by necessity or because they do not require additional staff and funding.

3. Establish implementation teams and a roster of advisors

For each of the prioritized roadmap elements, implementation teams will be assembled from across the campus based on expertise to develop and submit funding requests. The number of teams to be developed will depend on the number of roadmap elements approved by VCRCA Wakimoto and EVCP Hunt. These are NOT sinecure appointments; as appropriate, the leader of each implementation team will be expected to coordinate their peers to develop funding requests and create funding requests.¹⁰ At the same time, the coordinators

⁸ A search for 'astronomy' in the re3data.org registry of research data repositories returns 196 relevant results, some of which have already ceased operations (<https://www.re3data.org/search?query=astronomy>).

⁹ See Appendix 3 for a complete list of participants in the Phase 1 planning

¹⁰ The work required of the implementation teams will vary by roadmap element and could include writing job descriptions, conducting additional campus negotiations and discussions, updating quotes for equipment, determining funding allocation and staff effort, holding additional fact-finding forums, defining implementation strategies, and/or writing detailed specifications documents.

will pull together a roster of advisors to provide input to the implementation teams and administrative support as needed.¹¹

4. Initiate discussions with and submit requests to campus Budget & Finance

RDWG team members and advisors will begin conversations with UCLA's Budget & Finance team to alert them to the planned funding requests. As the budget requests are completed, the implementation coordinators and team leaders will work with Budget & Finance of their final form, review them with appropriate faculty and leadership groups, and present them to campus through the IT governance process.

5. Implement the funded roadmap solutions

As funding is allocated, implementation will proceed per the plan from the various teams and based on the information gathered for and included in the funding requests.

¹¹ This advisory group should include, but is not limited to, willing associate deans of research or equivalent to provide academic input, the IDRE board, IDRE and DataX leadership, faculty from relevant Academic Senate committees (e.g., CDITP, COR, COLASC), Library leadership, all facets of campus IT leadership (e.g., central, local, Health Sciences, OARC), APO, and Business & Finance.

Appendix 1: Meetings List

The following is a list of meetings and forums held to gather feedback on the RDWG roadmap and build support for the effort. Various members of the RDWG Steering Committee participated in the four faculty forums; RDWG co-chairs Todd Grappone and Lisa Snyder conducted the bulk of the meetings with faculty and academic leadership teams.

- Nov. 29: Committee on Data, Information Technology and Privacy (CDITP) (Academic Senate)
- March 1: Committee on Library & Scholarly Communication (COLASC) (Academic Senate)
- March 1: Council on Research (COR) (Academic Senate)
- March 7: EVCP Darnell Hunt
- April 5: Deans Council
- April 25: Institute for Digital Research and Education
- May 8: Roger Wakimoto (ORCA)
- May 15: Faculty Forum #1
- May 15: Annelie Rugg (Humanities)
- May 16: Eileen Stempel and Shannon McGarry (Music)
- May 16: Cindy Fan (International Institute)
- May 17: Faculty Forum #2
- May 22: Paul Krebsbach and Yvonne Kapila (Dentistry)
- May 22: Brian Kite, Jeff Burke, and Taylor Demma (Theater, Film and Television)
- May 23: Faculty Forum #3
- May 23: Tina Christie, Felipe Martinez, and Heath Hewitt (Education)
- May 23: UCLA Library Leadership
- May 23: Hilke Schlichting (Physical Sciences)
- May 24: Lucy Avetisyan (IT Services)
- May 26: Faculty Forum #4
- May 30: Abel Valenzuela, Kerri Johnson, and Mike Lee (Social Sciences)
- May 30: Aaron Burke (Faculty, Near Eastern Languages and Cultures)
- June 2: Organized Research Units (ORU) Directors (Betty Leung, Ruben Hernandez-Leon, Willeke Wendrich, Karen Burgess, Feli Schweizer, Arion Chatziioannou, Amy Rowat, Ali Behdad, Stephanie While, Stephen Cannon, and Veronica Terriquez)
- June 5: David Yoo (Institute of American Cultures)
- June 5: Ron Brookmeyer, James Macinko, and Jose Alvarez (Public Health)
- June 12: Physical Sciences Faculty Retreat (Miguel Garcia-Garibay, Rong Fu, Suzanne Paulson, Carolina Lithgow-Bertelloni, and a number of other unidentified participants)
- June 13: Anastasia Loukaitou-Sideris, Dennis Jong, Mits Yamahata (Public Affairs)
- June 13: Danny Snelson (Faculty, English)
- July 6: Dave Margolis, Jonathan Rogers, and Marcello Giannoni (Life Sciences IT leadership)
- July 6: Bruce Dunn, Greg Pottie, Todd Millstein (Engineering)
- July 17: Lin Zhan and Holli DeVon (Nursing)
- July 17: Antonio Bernardo and Margaret Shih (Anderson)

Aug 8: The Library Management Council

A meeting is pending with the academic leadership team in Life Sciences. The academic leadership teams of the David Geffen School of Medicine and the Law School did not respond to RDWG meeting requests.

Appendix 2: An RDWG Roadmap Implementation Task Analysis

To simplify analysis and prioritization, a spreadsheet was created with rows for each of the tasks called out in the RDWG roadmap or added after the discussion draft was generated in December of 2022. The text below explains the coding of the seven columns and is followed by the spreadsheet itself. For cost estimates, please refer to the original discussion draft released in December 2022.

Task Number

The task numbers reflect the numbering in the roadmap document. New items were added based on recommendations from the faculty forums and one-on-one interviews with the academic leadership teams. Rather than task numbers, these entries are identified in the spreadsheet by the source of the request.

Title/Task

A short description of the task.

Task Requirements

This column describes the actions that must be completed to make the task a reality (i.e., what needs to happen to make the task happen).

Status

Green

Indicates a task that could either move forward without significant funding given enough institutional will and a clear mandate for involved personnel or that seems to be moving forward as a result of other initiatives (e.g., the strategic plan, IT Services' campus technology overhaul).

Yellow

Indicates tasks that *might* be able to move forward without significant funding or that seem to be on the brink of moving forward (e.g., consultants are being interviewed for IT Services' network unification effort).

Red

Indicates tasks that cannot move forward without significant funding and/or dedicated personnel.

Strategic Plan

Indicates whether or not the task is explicitly included in the UCLA Strategic Plan in development. There are a handful of items marked YES, but most are UNKNOWN.

Priority Level

A column to enable prioritization. Tasks prioritized in the RDWG faculty forums or one-on-one meetings are marked as HIGH. All other tasks are left unprioritized.

Comments

Self-explanatory.

RDWG IMPLEMENTATION TASK ANALYSIS (version of July 12, 2023)

Task Number	Title/Task	Task Requirements	Status	Strategic Plan	Priority Level	Comments
RESEARCH DATA ROCKSTAR						
1	Hire an RDR or giving an existing person/unit the proposed responsibilities	Our Executive Sponsor group needs to make a decision to hire or otherwise assign these responsibilities to an existing employee. Decision needs to be made on administrative home. Funding needs to be allocated. If hiring, a hiring committee needs to be convened and a national search needs to happen.		Unknown	HIGH	
1	Hire additional team members to support the mission of the RDR.	Dependent on the hiring of the RDR or assignment of these responsibilities to an existing employee.		Unknown		
Supporting faculty and data-enabled research						
1.a	Continue planning for campus research infrastructure	Requires focused staff effort. Could be organized as standing committee. Would benefit from hiring of the RDR. Would require coordination with multiple departments on campus. Possible adoption of IDRE board as advisory entity.		Unknown		
1.b	Develop a faculty incentives program	Requires focused staff effort. Must be done in conjunction with the Office of the EVCP, ORCA, and the Academic Senate.		Unknown		
1.c.	Coordinate research facilitation	Could be lead by existing Research Concierge Program in coordination with existing research support personnel.		Unknown		
1.d.	Appoint a data-sharing compliance lead	May be the role under discussion in ORCA and RPC. At question is responsibility for P3/P4 data?		Unknown		
1.e	Represent UCLA at research data events and take leadership role in national organizations	Dependent on the hiring of the RDR or assignment of these responsibilities to an existing employee.		Unknown		
Based on Campus Feedback	Develop a program focused on proactive outreach to faculty regarding available resources, best practices, and compliance	Requires focused staff effort and a clear leader with official mandate.		Unknown		May be related to existing research concierge program
Developing faculty-focused education materials, guidance documents, and data-related training opportunities						
2.a	Create a cohesive research-related service catalog and web presence.	Requires focused staff effort and a clear leader with official mandate. May be part of the DSC and OARC missions.		YES	HIGH	Ideally owned by RDR; if none exists, where?
2.b	Create a funder and compliance requirements educational campaign	May be the role under discussion in ORCA and RPC.		Unknown		
2.c	Develop guidance on discipline- and domain-specific repositories	Requires focused staff effort and a clear leader with official mandate. May be part of the DSC mission.		YES		
2.d	Develop faculty outreach and training program on research-focused security, policy, and compliance	May be the role under discussion in ORCA and RPC.		Unknown		
2.e	Develop a data literacy effort	Requires focused staff effort and a clear leader with official mandate. May be part of the DSC mission.		YES		
Interface with faculty and campus leadership to improve institutional systems that negatively impact data-related research						
3.a.	Review processes that negatively impact research with an eye toward streamlining	Requires focused staff effort and cooperation from units that touch on the proposal process.		Unknown	HIGH	
3.b.	Work with campus stakeholders to create a long-term plan to restructure compliance-related efforts	Requires focused staff effort and cooperation from units that touch on the compliance process.		Unknown		
3.c.	Streamline processes related to compliance requirements	Requires focused staff effort and cooperation from units that touch on the compliance process.		Unknown		

RDWG IMPLEMENTATION TASK ANALYSIS (version of July 12, 2023)

Task Number	Title/Task	Task Requirements	Status	Strategic Plan	Priority Level	Comments
3.4.	Convene an Accountability Board	Requires focused staff effort and cooperation from units that touch on the compliance process.		Unknown		
RESEARCH DATA AND INFRASTRUCTURE SOLUTIONS						
STORAGE SOLUTION/TECHNOLOGY PLATFORM (aka BRUIN RESEARCH DATA HUB)						
1.1	Develop a hybrid on-premises/cloud-based research storage solution	Funding needs to be allocated and focused staff hired.		Unknown	HIGH	
1.2	Associated research data services to be built on top of the Bruin Research Data Hub					
1.2.a.	Data transfer nodes (DTNs)	Implementation dependent on the creation of the storage solution.		Unknown		
1.2.b.	Relocation of data currently stored in Google	Underway through IT Services. Data quotas for current faculty and staff still TBD.		Unknown		We have no plan for what happens next.
1.2.c.	License server	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Perhaps Software Central? Would require funding, dedicated staff effort, and coordination.		Unknown		If this item and the following items in this section were designed and implemented as part of the Bruin Research Data Hub, economies of scale could be realized.
1.2.d.	Research data commons (aka data mosh pit)	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Would require funding, dedicated staff effort, and coordination.		Unknown		Dependent on data transfer nodes?
1.2.e.	Database/dataset tools and service	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Would require funding, dedicated staff effort, and coordination.		Unknown		
1.2.f.	CoLab/JupyterHub/ELN environment	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Would require funding, dedicated staff effort, and coordination.		Unknown		
1.2.g.	Web hosting/sharing tools	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Would require funding, dedicated staff effort, and coordination.		Unknown		
Based on Campus Feedback	Virtual machines	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Would require funding, dedicated staff effort, and coordination.		Unknown		
Based on Campus Feedback	Containers	Could theoretically be constructed outside the Bruin Research Data Hub. The question is where? Would require funding, dedicated staff effort, and coordination.		Unknown		Containers can be used on Hoffman2
RESEARCH DATA SERVICES EXPANSION						
2.1.	Develop a Digital/Media Asset Management System (DAMS/MAMS)	Funding needs to be allocated and focused staff hired.		Unknown		A much-needed stand-alone system to be a long-term preservation solution for the campus.
2.2.	Expand consulting services and expertise					
2.2.a.	Expand the DSC to provide additional data-related training	Funding needs to be allocated and focused staff hired.		YES	HIGH	Some capacity exists, but simply not enough. DSC working with The Carpentries on data-related training courses.

RDWG IMPLEMENTATION TASK ANALYSIS (version of July 12, 2023)

Task Number	Title/Task	Task Requirements	Status	Strategic Plan	Priority Level	Comments
2.2.b.	Hire specialized innovation support personnel (e.g., data scientists)	Funding needs to be allocated and focused staff hired.		Unknown	HIGH	Some capacity for this type of consultation exists, but simply not enough.
2.3	Create a data services outreach and education program	Funding needs to be allocated and focused staff hired.		YES	HIGH	May be possible with existing staff but would require mandate shift
2.4.	Develop a program to digitize UCLA's primary source data	Funding needs to be allocated and focused staff hired.		Unknown		
RESEARCH INFRASTRUCTURE EXPANSION						
3.1.	Expansion of on-premises compute and storage					
3.1.a.	High-memory compute	Funding needs to be allocated. The Hoffman2 systems team would handle installation and administration.		Unknown	HIGH	
3.1.b.	GPU nodes	Funding needs to be allocated. The Hoffman2 systems team would handle installation and administration.		Unknown	HIGH	
3.1.c.	Compute nodes for general users	Funding needs to be allocated. The Hoffman2 systems team would handle installation and administration.		Unknown		
3.2.	P3/P4 compliant HPC/storage environment	A decision about how best to handle P3/P4 data is needed. Following that, funding must be allocated and focused staff hired. See also item 3.3.c.		Unknown	HIGH	
3.3.	Expand 'Cloud for Research' offerings					
3.3.a.	Campus cloud support for research	Decisions need to be made about the campus strategy for research use of cloud. If all in, funding must be allocated and focused staff hired.		Unknown		Some capacity exists, but simply not enough.
3.3.b.	Testbeds/Technology Sandbox	Funding needs to be allocated and focused staff hired.		Unknown		Some expertise exists within OARC, but not the band width.
3.3.c.	Possible cloud-based solution for P3/P4 data	Decisions need to be made about the campus strategy for research use of cloud. If all in, funding must be allocated and focused staff hired. See also item 3.2.		Unknown	HIGH	Also possible in on-premises solutions as noted above
	Additional service expansions					
4.1	Coordinate, expand, and promote UCLA's core facilities	Requires focused staff effort and cooperation from units that operate core facilities. See also item 2.a.		Unknown		The core facilities exist, but could be better integrated into educational and outreach campaign.
5.1	Identify and procure needed collaboration and research software tools	Requires focused staff effort, mandate, and funding.		Unknown	HIGH	
6.1	Identify and procure research data back-up solutions	Requires a decision as to how the campus will handle this, funding and agreement from the units that have a say in data backup solution (e.g., Security, Risk Management, ORCA, the units).		Unknown	HIGH	
7.1	Streamline research proposal development process	Requires focused staff effort and cooperation from units that touch on the proposal process.		Unknown		
7.1.a.	Establish a single source for guidelines, policies, and templates	Requires focused staff effort and cooperation from units that touch on the proposal process.		Unknown	HIGH	Connected to research support catalog noted above.
7.1.b.	Streamline workflow for proposal preparation	Requires focused staff effort and cooperation from units that touch on the proposal process.		Unknown		

RDWG IMPLEMENTATION TASK ANALYSIS (version of July 12, 2023)

Task Number	Title/Task	Task Requirements	Status	Strategic Plan	Priority Level	Comments
7.1.c.	Examine staffing related to the proposal development process	Requires focused staff effort and cooperation from units that touch on the proposal process.		Unknown		
7.1.c.i.	Develop resources to support research data compliance monitoring and support	Requires focused staff effort and cooperation from units that touch on the compliance process.		Unknown		
7.1.c.ii	Articulate and publish a responsibility matrix for data-related agreements	Requires focused staff effort and cooperation from units that touch on the compliance process.		Unknown		
7.1.d.	Identify and implement mechanisms for communicating data management and sharing plan requirements	Requires focused staff effort and cooperation from units that touch on the proposal process.		Unknown		
Implicit	Develop a security and compliance program integrated with research	Requires focused staff effort and cooperation from units that touch on the research process.		Unknown	HIGH	
Implicit	Streamline interactions between the campus and the health sciences	Requires focused staff effort and cooperation from units that touch on the research process.		Unknown		
Based on Campus Feedback	Identify and implement co-location options	Requires physical space, administrative infrastructure (e.g., recharge mechanism), funding for facilities upgrades		Unknown		Having co-lo options for researchers will minimize the number of systems across campus
Based on Campus Feedback	Identify and implement hosting options	Requires physical space, people to support and maintain the equipment, administrative infrastructure (e.g., recharge mechanism), funding for facilities upgrades		Unknown		Having hosting options for researchers will minimize the number of systems across campus
RECURRING REQUESTS FROM ACADEMIC LEADERSHIP TEAMS						
Based on Campus Feedback	Research business analytics	Requires funding and focused staff effort.		Unknown		Has been discussed over the years to no avail.
Based on Campus Feedback	Faculty profiles	Requires funding and focused staff effort.		Unknown		Has been discussed over the years to no avail.
KNOWN ELEMENTS FROM STRATEGIC PLAN						
Strategic Plan	Team to assist with large-scale research proposals	Planning underway within ORCA		YES		To be confirmed with Mary Okino in ORCA
Strategic Plan	Improve proposal tracking and other grant reporting requirements.	Already underway, ongoing effort		YES		Incremental improvements have been made over the last 10 years. Proposal tracking is done in PATS, and visible to campus via ORA Portal, PI Portal
Strategic Plan	Add additional search capabilities in PATS	Extensive proposal and award search capabilities of PATS data are already available in ORA Portal		YES		PATS is a tool used by central administrators. Extensive search capabilities of PATS data are already available in ORA Portal.
CURRENTLY ON TO-DO LIST FOR IT SERVICES						
Underway (IT Services)	Network unification	Currently a pending project with no known completion date		Unknown	HIGH	Completion is required to enable certain roadmap elements
Underway (IT Services)	Modern identity and access management system	Currently listed as a key initiative under information security with an assigned project manager with no known completion date		Unknown	HIGH	Completion is required to enable certain roadmap elements

Appendix 3: Participants in Phase 1 of the RDWG Effort

Executive Sponsors

Ellen Pollack, Chief Information Officer, UCLA Health Sciences; Chief Nursing Informatics Officer
Jim Davis, Vice Provost IT and Chief Academic Technology Officer
Lucy Avetisyan, Associate Vice Chancellor and CIO
Roger Wakimoto, Vice Chancellor for Research and Creative Activities
Virginia Steel, Norman and Armena Powell University Librarian

Steering Committee

Lisa M. Snyder, Director, Campus Research Initiatives/Advanced Research Computing Research Technology Group, Office of Advanced Research Computing (co-chair)
Todd Grappone, Associate University Librarian for Digital Initiatives and Information Technology, the UCLA Library (co-chair)

Albert Duntugan, Chief Data Officer, UCLA Health Sciences
Alex Bui, Professor, Radiological Sciences, Bioengineering, and Bioinformatics; and Director of the UCLA Medical & Imaging Informatics (MII) group; and chair of the Academic Senate Committee on Data, Information Technology and Privacy
Bill Holloway, Chief Academic Information Officer, UCLA Health Sciences
Deanna Needell, Professor, Mathematics, and Executive Director of the Institute for Digital Research and Education
George Morris, Research Infrastructure Consultant to the Office of Advanced Research Computing
Howard Miller, CIO, Anderson School of Management
Jackson Jeng, Senior Director, Office for Research Information Systems, Office of Research Administration
Jacob Foster, Associate Professor, Sociology, and co-lead of the DataX Initiative
Marcia Smith, Associate Vice Chancellor UCLA Research Administration
Mary Okino, Deputy Director, Office of the Vice Chancellor for Research and Creative Activities
Michael Van Norman, Chief Technology Officer, IT Services
Nico Wibowo, IT Director, SEASnet, Samueli School of Engineering
Sarah Klein, Pre-Award Coordinator, Semel Institute
Tim Dennis, Director, Data Science Center, the UCLA Library

Technical Infrastructure Subcommittee

Michael Van Norman (chair)

Brian Pape, Manager, Hoffman2 Systems Research Technology Group, Office of Advanced Research Computing

Edson Smith, Data Management Strategist, UCLA Library

George Morris

Lisa M. Snyder

Mike Shane, Senior Director, Platform Services, UCLA Health Sciences

Nico Wibowo, Director, SEASnet, Samueli School of Engineering

Tito Deveyra, Director, Information Technology, Theater, Film and Television (retired)

TV Singh, Chief Computational Scientist, Office of Advanced Research Computing

Data Services Subcommittee

Tim Dennis (chair)

Clifford Kravit, Program Manager, Research Enablement, Health Sciences

Eleanor Koehl, Research Facilitation Coordinator, Office of Advanced Research Computing

Ibraheem Ali, User Engagement Division, UCLA Library

Keith Chen, Professor, Anderson School of Management

Rose Rocchio, Director, Mobile and Web Strategy, Office of Advanced Research Computing

Todd Grappone

Yoh Kawano, GIS and Visualization Manager, Office of Advanced Research Computing

Cloud Vendors and Services Subcommittee

Bill Holloway (co-chair)

George Morris (co-chair)

Andrew Browning, Manager, Research Web and Data Platforms, Office of Advanced Research Computing

Clifford Kravit

Hector Felix, Director, Enterprise Technology, IT Services

Lisa M. Snyder

Mike Shane

Mike Van Norman

Shao-Ching Huang, Manager, Computational Science Research Technology Group, Office of Advanced Research Computing

Shehzad Sheikh, Director, Enterprise Data and Analytics Platform, UCLA Health Sciences

Research Proposal Process and Data Subcommittee

Sarah Klein, Pre-Award Coordinator, Semel Institute (chair)

Jackson Jeng, Senior Director, Office for Research Information Systems, Office of Research Administration

Lisa M. Snyder

Marcia Smith

Mary Okino

Todd Grappone

Proposed and Pending Subcommittees

The **Security, Privacy, Policy, and Compliance Subcommittee** was intended to analyze the completeness of existing and the implications of all known and proposed security, privacy, policy, accessibility, and compliance requirements (e.g., HIPAA, PCI, FERPA, CMMC for DoD research) and propose solutions to address those requirements and confirm secure computing solutions and workflows that support data use, risk, manipulation, and publishing. This subcommittee was never activated and it is anticipated that it will be convened at the point of funding and implementation.

The **Faculty Incentives Subcommittee** was intended to work with the Academic Senate, the Academic Personnel office, and faculty representatives to determine and implement incentives for best practices related to research data and open data. This subcommittee was never activated and it is anticipated that it will be convened at the point of funding and implementation.

The **Public Access to Research Data Subcommittee** was intended to ensure that UCLA data are made openly available through the most appropriate means possible (e.g., data publishing, data sharing, discipline-specific repositories). This subcommittee was never activated and it is anticipated that it will be convened at the point of funding and implementation.

The **Research and Teaching Interface Subcommittee** was intended to ensure the coordination and interoperability of research and research data with Canvas, UCLA's primary learning management system, and use of UCLA data in labs and other training programs. This subcommittee was never activated and it is anticipated that it will be convened at the point of funding and implementation.

The **Data Diversity and Ethical Use Subcommittee** was intended to review algorithmic data use to ensure quality and fairness, discussions of bias, ethics, and provenance. This subcommittee was never activated and it is anticipated that it will be convened at the point of funding and implementation.