Research Data Working Group (2022): Findings and Recommendations for OneIT

Reviewed by Tim Dennis, UCLA Library Data Science Center

2026-03-30

Source: *A Call to Action: A Roadmap for Expanding Research Infrastructure and Data Services at UCLA* (December 2, 2022). Commissioned by five executive sponsors: VCR, University Librarian, AVC & CIO, Vice Provost for IT, and CIO of Health Sciences. 18-month effort; campus-wide engagement and peer benchmarking across 13 service categories.

Key Findings That Remain Relevant Today

- **UCLA falls behind peers in 8 of 13 research data service categories.** The lowest score was for a cohesive research data services program with a website and service catalog — the most basic coordination infrastructure.
- **90% of UCLA's $7M annual enterprise storage spend occurs at the departmental level.** This fragmentation produces redundancy, inconsistent security, and no economies of scale.
- **There is no central ownership of research data services.** Multiple units respond to campus-wide issues (funder mandates, storage crises, compliance requirements) independently and inconsistently. PIs navigate this alone: *"death by a thousand cuts."*
- **UCLA's NSF per-award average ($307K) is 15–30% below UCSB ($428K) and UCSD ($412K).** The RDWG identified missing infrastructure for large, multi-site proposals as a contributing factor.
- **Funder mandates have outpaced UCLA's capacity.** The NIH DMSP (effective January 2023) and the OSTP public access mandate (effective December 2025) require institutional infrastructure UCLA does not yet have in place.

Relevant Recommendations for OneIT

The RDWG recommended a **layered model** — not consolidation, but coordination. Central services raise the floor; local expertise is preserved where research happens. This maps directly onto what OneIT can address:

1. Create a central research data coordination function. The RDWG proposed an Office of the Research Data Rockstar (ORDR) to consolidate ownership, accountability, and faculty support. Three years later this role still does not exist. OneIT's governance structure is the right place to establish it — with authority under the VCR, not IT operations.

2. Establish a shared research storage platform with baseline faculty allocation. The proposed Bruin Research Data Hub would provide every tenure-track faculty member a baseline storage allocation at no cost. This directly addresses the 90% departmental fragmentation finding and is the highest-leverage infrastructure investment the RDWG identified.

3. Centralize what scales; preserve what requires domain expertise. Networking, IAM/SSO, cybersecurity, cloud procurement, and shared storage are appropriate for central consolidation. Research computing, data services consulting, and domain-specific systems (HPC, secure enclaves, discipline-specific repositories) require embedded expertise and researcher proximity that does not survive enterprise IT consolidation. Peer institutions that have outpaced UCLA share this model.

4. Build the P3/P4 compliant environment for restricted data. Export controls on AI and semiconductor research, clinical data requirements, and ITAR-adjacent research are a growing share of UCLA's portfolio. This infrastructure was recommended in 2022 and is now more urgent.

Evidence Supporting Research Enablement Needs

- **Peer benchmarking (13 categories):** Available in the RDWG report and *Research on Data Services II.* Identifies specific gaps relative to Michigan, Stanford, UCSD, and others — most of which have invested in coordinated research data infrastructure since 2022.
- **NSF and NIH award data:** UCLA's per-award gap and the scale of its NIH portfolio ($505M annually) document both the opportunity and the compliance risk.
- **Researcher use cases:** Appendix 3 of the roadmap includes discipline-specific scenarios documenting where infrastructure gaps directly impede research.
- **Cost of inaction:** The Google Drive 10PB migration crisis (June 2024 deadline), estimated $1.8M/year in potential overage fees, is a concrete recent example of what distributed, uncoordinated storage costs.
- **Phase 2 materials (2023–2024):** Implementation task analysis, Deans Council brief, and Bruin Data Hub specifications provide additional detail on feasibility and sequencing.

All source documents available for download on the Resources page.

What Has Changed Since 2022 — and What Persists

What has changed:

- **NIH DMSP took effect January 2023** — six weeks after the roadmap published. Compliance is now required, not anticipated.
- **OSTP public access mandate took effect December 2025** — all federally-funded research must be publicly available at no cost. This is present reality, not a future risk.
- **Generative AI** has created cross-disciplinary GPU compute demand the roadmap did not anticipate. Access to AI infrastructure is now a research equity issue on par with storage in 2022.
- **Federal funding uncertainty has intensified** (March 2026): UCLA's ability to compete for large, multi-site awards is more consequential than ever.
- **One IT launched**, raising new organizational questions about research IT roles that the RDWG did not address.

What no longer applies as written:

- Cost figures ($74.2M / $49.8M) reflect 2022 market conditions; updated estimates would be needed.
- "Future" mandates (NIH DMSP, OSTP) are now in force — the urgency framing has shifted from anticipatory to reactive.

What persists:

- UCLA's benchmarking gaps remain: none of the four core recommendations have been implemented.
- The coordination vacuum (no ORDR equivalent) is still unaddressed.
- Faculty reliance on informal workarounds — servers under desks, fragmented storage, ad hoc cloud spend — continues to be documented in recent campus discussions. It matches exactly what the RDWG found in 2022.
- **The GIS librarian vacancy is a concrete local example of the mechanism in action.** A December 2025 report by OARC, Software Central, and the Library on ArcGIS Online governance documented the positive arc clearly: a GIS specialist built out the platform, broadened access, and drove campus adoption — usage grew. When that position became vacant, the platform that researchers were actively depending on lost its designated administrator and governance structure. It now operates at 230-240% of licensed storage capacity with no retention policy and real risk of Esri service disruption. (Horne & Woltag, OARC/Software Central/Library, December 2025)
- The layered model remains the right architecture: **central coordination raises the floor; embedded expertise stays where research happens.**