https://images.ucla.edu/bp/#/folder/173413/atkinson-01.00406-royce-shapiro

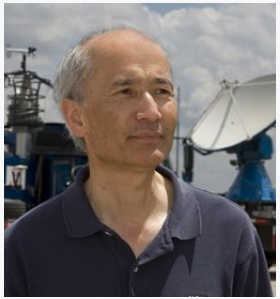
UCLA

RESEARCH DATA ROADMAP OVERVIEW

WINTER QUARTER 2023

EXECUTIVE SPONSORS

In early 2021, five UCLA administrators charged a Research Data Working Group to develop a high-level roadmap for research data services and infrastructure at UCLA.



ROGER WAKIMOTO

Vice Chancellor for Research & Creative Activities

GINNY STEEL

Norman and Armena Powell University Librarian

LUCY AVETISYAN

Associate Vice Chancellor and CIO

JIM DAVIS

Vice Provost IT and Chief Academic Technology Officer

ELLEN POLLACK

CIO, UCLA Health Sciences; Chief Nursing Informatics Officer

RDWG STEERING COMMITTEE

Alex Bui, Professor, Radiological Sciences, Bioengineering, and Bioinformatics; Chair, Academic Senate Committee on Data, Information Technology and Privacy
Deanna Needell, Professor, Mathematics; Executive Director, Institute for Digital Research and Education
Jacob Foster, Associate Professor, Sociology; Co-Lead, DataX Initiative

Lisa M. Snyder (Co-Chair), Director, Campus Research Initiatives/Advanced Research Computing Research Technology Group
George Morris, Research Infrastructure Consultant

Todd Grappone (Co-Chair), Associate University Librarian for Digital Initiatives and Information Technology
Tim Dennis, Director, Data Science Center, the UCLA Library

Bill Holloway, Chief Academic Information Officer, UCLA Health Sciences
Albert Duntugan, Chief Data Officer, UCLA Health Sciences

Mary Okino, Deputy Director, Office of the Vice Chancellor for Research and Creative Activities

Marcia Smith, Associate Vice Chancellor UCLA Research Administration
Jackson Jeng, Senior Director, Office for Research Information Systems, Office of Research Administration
Sarah Klein, Pre-Award Coordinator, Semel Institute

Mike Van Norman, Chief Technology Officer, IT Services

Howard Miller, CIO, Anderson School of Management
Nico Wibowo, IT Director, SEASnet, Samueli School of Engineering

UCLA

FOUR RDWG SUBCOMMITTEES

Four RDWG subcommittees were convened with five others pending activation.

TECHNICAL INFRASTRUCTURE

Chair: Mike Van Norman

Charged with defining the technical elements that need to be in place for research data support (i.e., identifying the 'should be' research environment for UCLA).

CLOUD

Chair: Bill Holloway

Focused on UCLA's current and future use of cloud services and recommendations for on-premises and cloud options that should be available to researchers.

SECURITY, PRIVACY, AND COMPLIANCE

Chair: Albert Duntugan

Pending activation. Charged with analyzing all known and proposed security, privacy, policy, accessibility, and compliance requirements and ensuring that they are addressed in the RDWG roadmap.

DATA SERVICES

Chair: Tim Dennis

Charged with defining the necessary suite of services and software needed to support the research/data lifecycle campuswide.

GRANT PROPOSAL PROCESS

Chair: Sarah Klein

Charged with reviewing existing campus practices for research proposals and identifying process points that intersect with planned research data support services.

PLANNED SUBCOMMITTEES

As implementation for the RDWG roadmap begins, four additional subcommittees will be activated: Faculty Incentives, Public Access to Research Data, the Research and Teaching Interface, and Data Diversity and Ethics.

UCLA

DEVELOPMENT OF THE RDWG ROADMAP

Over two years, the RDWG Steering Committee and subcommittees met to discuss research services and infrastructure at UCLA. The resultant roadmap was informed by four key sources of information:

1 PAST UCLA RESEARCH INFRASTRUCTURE EFFORTS

The RDWG roadmap builds on past UCLA infrastructure planning efforts including the 2013 North Campus Infrastructure Planning Summit; the first phase of the 2013-2014 Research Informatics Strategic Planning (RISP) Initiative; and the 2017 Research Data Support Recommendations.

2 RESEARCH SUPPORT OFFERINGS AT PEER INSTITUTIONS

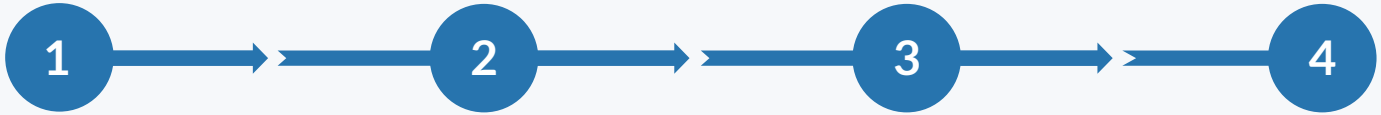
Peer institutions have outpaced UCLA with infrastructure and realignment around research data. While we at UCLA have been talking about research data, other campuses have moved forward with investments and an expansion of their support offerings, with UCLA falling considerably behind our peers in benchmarking studies.

3 DISCUSSIONS WITH FACULTY AND RESEARCH SUPPORT PERSONNEL

The RDWG has met and evaluated data and experiences from a broad cross-section of the research community. In addition, representation from the major IT partners across UCLA joined in the discussion and analysis.

4 EVOLVING REQUIREMENTS FROM MAJOR FUNDERS

Our major federal and nonprofit sponsors now have mandates that require data sharing in order to accelerate research, encourage multi-institutional collaboration. Policies and enforcement plans for mandatory data sharing that promote open research are now in place to ensure that research funded by taxpayers and donors is of optimal benefit to the public.



UCLA

PROBLEMS & ISSUES

The fundamental problems are relatively straightforward, but any discussions about possible solutions need to address the many confounding issues that have hobbled past efforts to develop campuswide research data services and infrastructure.

SHIFT IN RESEARCH CULTURE

UCLA is not prepared to support faculty in our new era of scholarship that prioritizes interdisciplinarity, collaboration, open science, open data, and new requirements for security, privacy, and compliance.

DATA-SHARING REQUIREMENTS

Changes are required to meet the institutional responsibility for compliance with the policies federal funding agencies have set for mandatory data sharing to promote open research.

FACULTY NEEDS

An imperative for UCLA is to ensure that the needs of faculty and researchers in all units are met by a blend of local and central campus infrastructure and services.

DECENTRALIZATION

The result of a funding model built on subsidiarity is that support for faculty varies widely across campus with isolated systems that limit collaboration and transparency.

DEFERRED COMMITMENT AND INVESTMENT

The longer we wait to address these challenges, the harder it will be to build the necessary services and resources.

POTENTIAL FOR UC/AGENCY ACTION

Promised services and the potential for outside funding have prevented action at the scale needed at UCLA.

COORDINATION AND ORGANIZATION

Ownership, authority, and responsibility related to research data are still outstanding questions for discussion.

ROADMAP EXECUTIVE SUMMARY

To ensure excellence in the future, UCLA needs to make a standing commitment to invest in research data services and infrastructure, and we need to take action immediately. Providing faculty with access to fundamental research infrastructure and data services that support their ability to generate, store, manipulate, and analyze data is an institutional imperative if UCLA wants to continue to lead in this new era of open research. Investing in central services enhances distributed efforts, supports compliance, and simplifies collaboration.

STATE OF THE RDWG ROADMAP

The RDWG's draft roadmap is a great starting point for conversations with the campus about the research data services and infrastructure needed to ensure UCLA's future success.

FACULTY ALIGNMENT

The RDWG is embarking on a series of meetings to ensure that the roadmap recommendations are aligned with faculty needs and priorities, and, if not, gather the information required to revise the document as necessary.

ADMINISTRATIVE BUY-IN

Simultaneously, the RDWG is working to build consensus at the administrative level on the desirability of centralized research data services and infrastructure in support of distributed data-intensive research and partnerships.

UCLA

OUR VISION

The key elements of our vision for research data services and infrastructure are centered on a full lifecycle approach to open science and scholarship that will position UCLA faculty and researchers to continue to be competitive for larger, multi-institutional proposals built on the principles of sharing, collaboration, training, and best practices.

BLENDED SUPPORT MODEL

Research support that combines discipline-specific resources at the local level with central baseline offerings and services.

GUIDANCE FOR RESEARCHERS

Coherent and cohesive information about available research data services and resources, and policy requirements and practices.

BASELINE INFRASTRUCTURE

Baseline research data storage, management, computation, and access through infrastructure available at no cost to all faculty.

STREAMLINED ACCESS

Clear pathways to resources and services for faculty with appointments both on campus and in the Health Sciences.

ACCOUNTABILITY

Clear administrative responsibility for research infrastructure and data services aligned with faculty governance.

DATA PROTECTIONS

Security, privacy and accessibility integrated with research needs and focused on protecting UC's most valuable data assets and IP.

RESEARCH DATA PRINCIPLES

The RDWG roadmap hinges on a basic set of principles related to the research data lifecycle.

OPEN ACCESS

UCLA is committed to ensuring that all "published research articles in the broadest sense of the term" authored by any UC employee be made available to the public at no charge.

ADHERENCE TO 'FAIR' PRINCIPLES

UCLA expects that this mandate be extended to include other forms of research outputs in accordance with the FAIR guiding principles.

SUPPORTING DATA SHARING

UCLA is also committed to support researchers in meeting data-sharing goals and developing the infrastructure necessary to support best data practices across the research lifecycle.

DATA MANAGEMENT & CURATION SERVICES

Critical to fulfillment of this commitment is the development of a data management, preservation, and curation services

NEED FOR DATA INFRASTRUCTURE

Infrastructure for storing and preserving UCLA's digital assets is also critical and should be prioritized by the campus

FREE DATA SERVICES

UCLA is committed to providing data services and infrastructure to researchers at no cost or at a substantially subsidized rate.

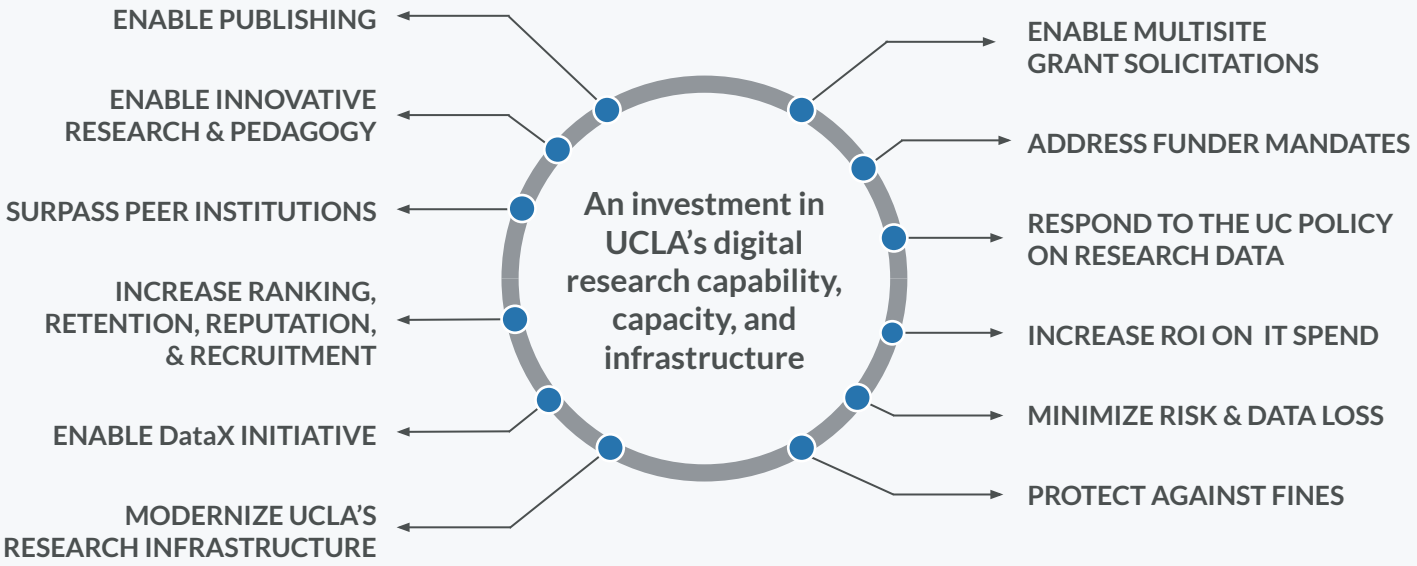
DATA PRESERVATION PRIORITIZATION

Conversations are needed at the highest levels to prioritize research outputs for curation and preservation.

GUIDANCE AND INCENTIVES

The campus needs to develop incentives to encourage data best practices, clear guidelines for researchers, and a transparent and accountable governance structure.

KEY BENEFITS AND OPPORTUNITIES



https://www.uclahealth.org/news/38-ucla-scientists-among-worlds-most-influential-researchers

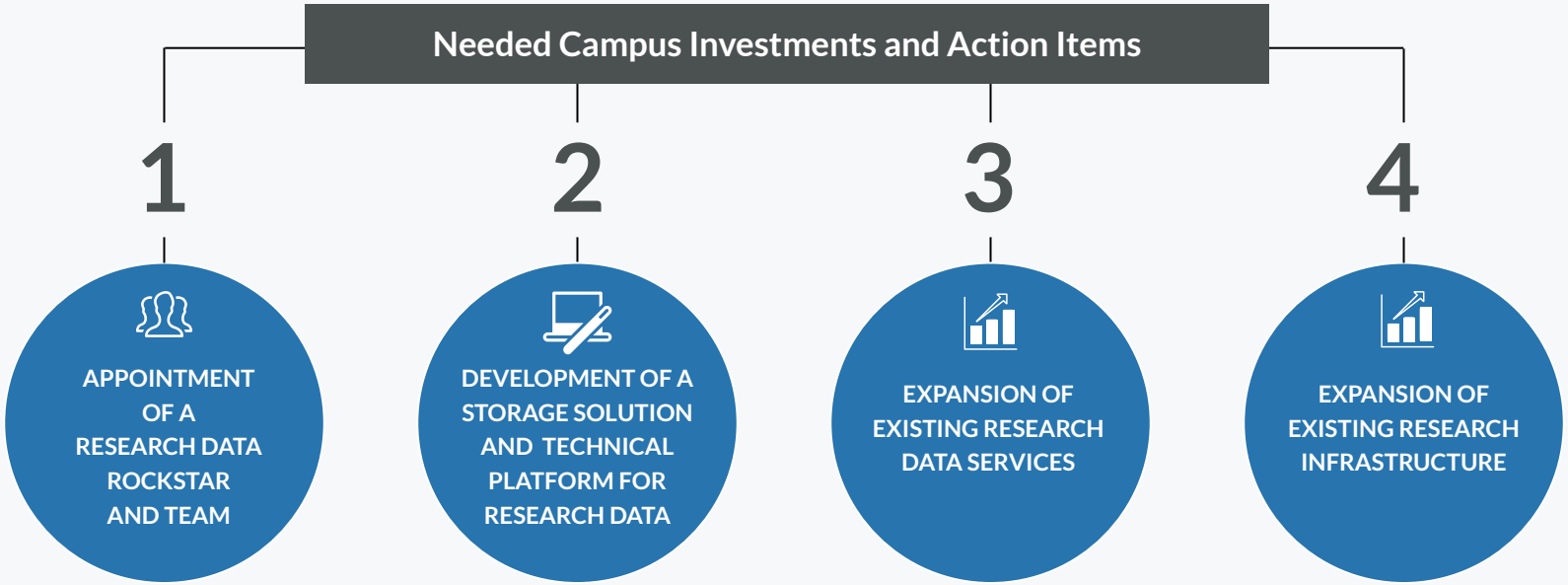
UCLA

MAIN ROADMAP ELEMENTS

UCLA

RDWG ROADMAP RECOMMENDATIONS

The RDWG roadmap calls for campus investments in four key research data services and infrastructure areas.



https://images.ucla.edu/bp/#/folder/173370/jsc2017e067121 jessica-watkins-nasa

UCLA

RESEARCH DATA ROCKSTAR

UCLA

RESEARCH DATA ROCKSTAR (& TEAM)

The Research Data Working Group's first assertion is that the campus needs to formally consolidate and clarify ownership, responsibility, and accountability for initiatives related to research infrastructure and research data services. One of the key challenges of our currently decentralized research support network is that there is no clarity about ownership of issues or situations with campuswide implications.

SUPPORT FACULTY AND DATA-ENABLED RESEARCH

- Continue planning for expansion of campus research data services and infrastructure
- Develop faculty incentives program
- Coordinate research facilitation
- Appoint a data-sharing compliance lead/officer
- Represent UCLA in data-related national/international organizations

FACULTY-FOCUSED MATERIALS AND GUIDES

- Create a cohesive research-related service catalog and web presence
- Create a funder and compliance requirements educational campaign
- Develop guidance on discipline- and domain-specific repositories
- Develop data- and security-related outreach and training programs

TRANSPARENCY AND ACCOUNTABILITY

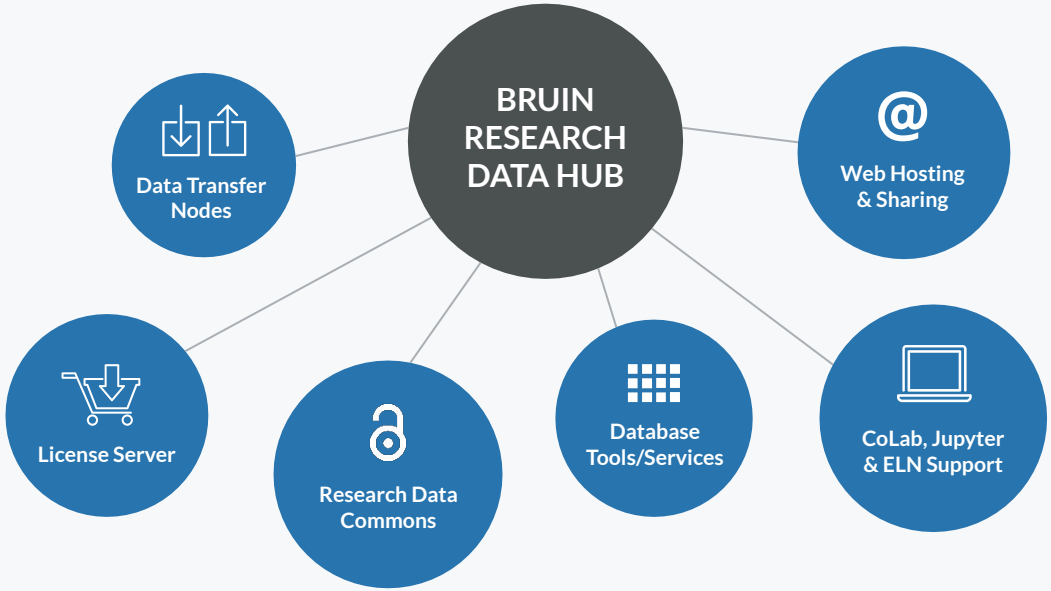
- Review and streamline processes that negatively impact research
- Create a long-term plan to restructure compliance-related efforts
- Streamline compliance requirements
- Convene Accountability Board

UCLA

STORAGE/TECHNOLOGY SOLUTION

STORAGE/TECHNOLOGY SOLUTION

The RDWG recommends equipping the campus with a technology platform that combines high-performance storage and software to create a hybrid on-premises and cloud solution that is scalable to meet current and future research data needs. As envisioned, this platform will be the home for campuswide research data services, include tools to facilitate interdisciplinary work, and provide storage space for members of the UCLA research community with shareable lab directories for PIs.



https://images.ucla.edu/bp/#/folder/173427/130514_UCLA_9579-cropped

UCLA

EXPAND EXISTING RESEARCH DATA SERVICES

EXPAND RESEARCH DATA SERVICES

Expanding and streamlining existing research data services is necessary to respond to funder and scholar mandates and expectations. While many of these services have been recently piloted at UCLA, they will need to be expanded and fully funded to meet expectations as outlined in this roadmap.

EXISTING DATA SCIENCE CENTER

Provides research support focusing on promoting openness, transparency, and reuse for data and software used in a publication.

DATA SERVICES OUTREACH & EDUCATION PROGRAM

Development of a training program for local support that advocates best practices in their domains, the repositories chosen for their work, and sociotechnical issues related to data use.

EXPANSION OF STAFF EXPERTISE TO SUPPORT RESEARCH

The additional work required to support faculty and address the complexity of ever-changing mandates and compliance challenges will require new dedicated staff.

EXISTING DIGITAL LIBRARY PROGRAM

The largest digital library program in the UC system offers digitization and preservation services locally as well as for dozens of international partners.

DIGITAL/MEDIA ASSET MANAGEMENT

A system where researchers deposit research objects and metadata that make them discoverable, findable, downloadable, storable, and interoperable within the larger scholarly communication ecosystem (publication systems, data repositories).

DIGITIZATION OF PRIMARY SOURCE DATA

Digitizing, describing, and processing Library collections so that they are accessible research materials for digital scholarship (e.g., text mining, computer vision, OCR structured data).

https://images.ucla.edu/bp/#/folder/202734/ucla_day4_UCOP-1

UCLA

EXPAND EXISTING RESEARCH INFRASTRUCTURE

EXPAND RESEARCH INFRASTRUCTURE

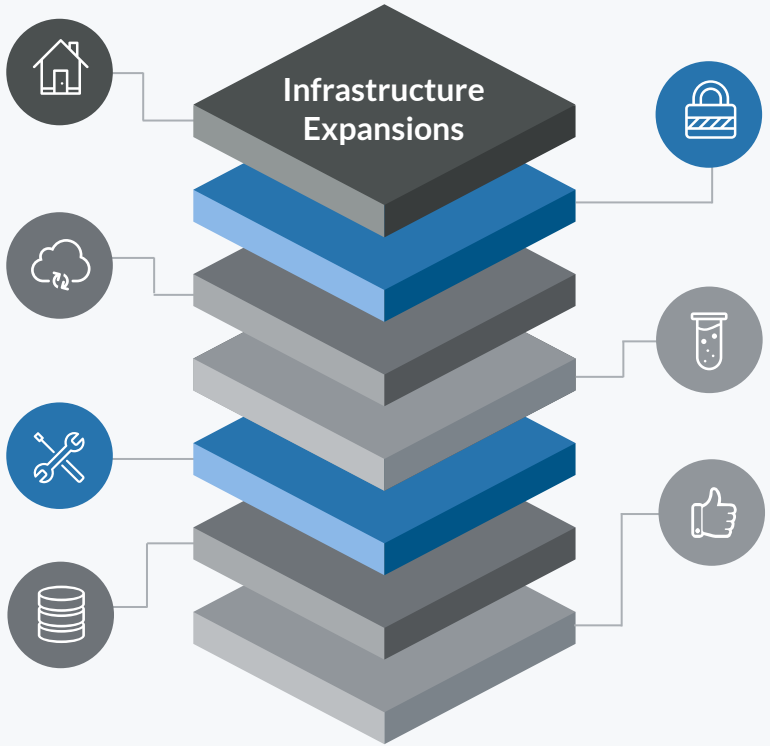
In addition to the development of a new storage and technology platform, supporting research at UCLA will also require a commitment to expanded investment in existing infrastructure and the development of much-needed services that address specific research support gaps.

EXPANSION OF ON-PREMISES COMPUTE & STORAGE
- High-memory compute
- GPU nodes
- Compute nodes for general users

EXPANSION OF 'CLOUD FOR RESEARCH' OFFERINGS
- Cloud support for research
- Testbeds/Technology Sandbox
- Alternate P3/P4 environment

COLLABORATION AND RESEARCH SOFTWARE AND TOOLS

RESEARCH DATA BACKUP



P3/P4 COMPLIANT HPC/STORAGE ENVIRONMENT

COORDINATION, EXPANSION, AND PROMOTION OF THE CORE TECHNOLOGY CENTERS THAT SUPPORT UCLA RESEARCH

STREAMLINING OF PROPOSAL PROCESS
- Single source of truth for data guidelines
- Streamline workflows
- Proposal development staffing
- DMSP requirements updates

https://images.ucla.edu/app/#/folder/173426/130515_UCLA_3238

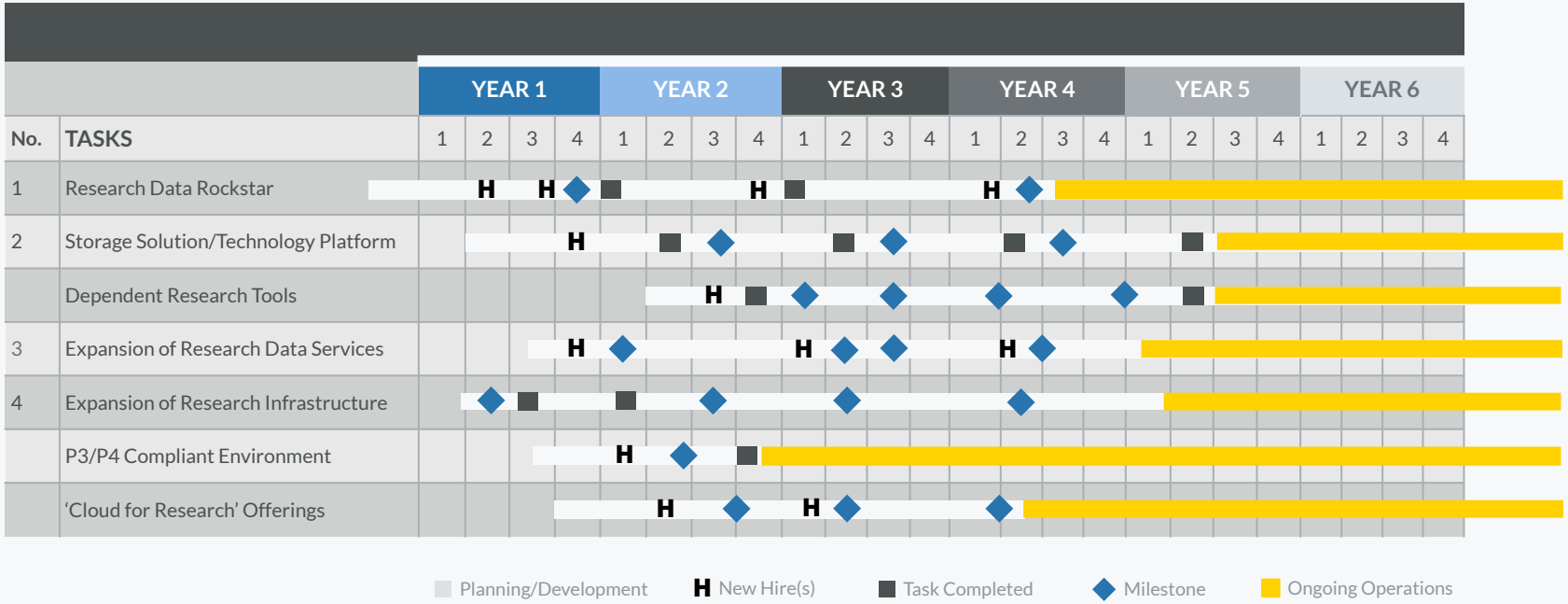
UCLA

IMPLEMENTATION

UCLA

PROJECT TIMELINE IF FULLY IMPLEMENTED

From the moment of authorization, full implementation of the roadmap would require concurrent development of each of the major elements.



“

UCLA HAS EXPERIENCED AN ASTONISHING ASCENT IN ITS FIRST HUNDRED YEARS, AND IN OUR SECOND CENTURY WE CONTINUE TO RISE, MAKING AN EVEN GREATER IMPACT UPON THE COMMUNITIES WE SERVE.

”

GENE D. BLOCK

https://images.ucla.edu/bp/#/folder/173515/ucla_day4_UCOP-27

UCLA

RESEARCH DATA ROADMAP DISCUSSION

WINTER QUARTER 2023