Blueprint 2030

The Digital University in the Next Decade

David Minor
Director, Research Data Curation Program
Library

Valerie E. Polichar
Sr. Director, Academic Technology Services
IT Services

Co-Chairs,
UC San Diego Research Computing & Data Services Governance Committee

in consultation with UC San Diego administration & faculty

January 2022

Blueprint 2030: origins

In 2009, the original *Blueprint for the Digital University* set the stage for UC San Diego's many accomplishments in cyberinfrastructure over the following decade. Many contemporary services and infrastructure used across campus have their origins in this document and subsequent projects. But research evolves, new needs arise, and new requirements are placed on researchers. Thus, campus leadership recognized the need for a new, updated Blueprint, helping point the way to the next generation of support.

To compile this new Blueprint, the Research Computing and Data Services Governance Committee engaged in a multi-year process to obtain expert input into the future of research and research technology, including consulting a faculty advisory group, surveying campus faculty, engaging in in-depth interviews with UC San Diego administrators such as the Executive Vice Chancellor and Vice Chancellor for Research, and conducting faculty focus groups.

This document outlines the priorities and directions our University must take to maximize new discoveries, increase funding, facilitate science, comply with regulatory demands and propel us forward in the decade to come.

The themes described in this document come from trends in research processes and workflow, and are not specific to UC San Diego. This document is being shared publicly with the hope that it will be of assistance to other institutions facing similar generational evolution in both the conduct of research and in their research support infrastructure.

The Blueprint 2030 project, presented here:

- represents the outcome of two+ years of discussions with campus leadership and researchers on anticipated developments in the research space, and the status of today's research technology support;
- coalesces these discussions into themes that were raised across campus and disciplines;
- based on these discussions, lays out strategic priorities to enable our campus to continue to lead the world in research;
- identifies current research technology strengths and challenges, and how they could be addressed in the future;
- makes no specific assumptions on future technology or staffing choices.

This document is the first step in a cross-campus research technology design process that will ensure that UC San Diego facilities keep pace with contemporary research needs. Next steps include:

- determining and prioritizing concrete technologies and services in response to the themes presented in this document;
- subsequent gap analyses between current and goal technology infrastructure/services and the desired future priorities;
- discussion of budget implications and appropriate funding opportunities.

UC San Diego in 2030: a home for vibrant, collaborative research

In 2030, multidisciplinary and multi-institutional partnerships will be the norm, with teams of researchers working together seamlessly, sharing large datasets and integrating their workflows smoothly no matter their discipline. UC San Diego researchers will be closely integrated with local city and county governments, working together on projects that benefit the region and stimulate student interest.

UC San Diego will remain a center for cutting-edge research on technology, advancing the state of the art in computing, telecommunications, robotics, bioengineering and other disciplines, informed by knowledge from the arts, humanities and social sciences. Researchers will leverage multiple campus core services to use high-end instruments and analysis services producing immense amounts of data to increase scientific knowledge and solve real-world problems. Novel research products like source code, creative works, and data itself will garner acclaim for our faculty. Our researchers will easily tap into huge shared datasets (such as government, climate, or social media) to produce analyses and to train scientific models. They will employ remote instrumentation and sensor networks to research the oceans, the surface of the planet, and outer space. This wide range of new technologies and services, such as reproduction and study of artifacts, text, and works of art, and the ability to run large-scale analysis of data collections, will also revolutionize scholarship in the arts, humanities and social sciences. And UC San Diego researchers and students will participate regularly in remote experiences from the comfort and beauty of UC San Diego's campus — traveling seagoing vessels, exploring research sites worldwide, and interviewing human subjects thousands of miles away.

This is the vision for UC San Diego shared with us by our faculty members and administrators; this trajectory is already visible in current trends.

Contents

Executive Summary

UC San Diego's past investments in research cyberinfrastructure made possible enormous accomplishments; but neither research nor technology sits still. In 2020, campus administration charged the Research Computing and Data Services Governance Committee to begin the process of creating a new long-term strategic plan for research support. This document, *Blueprint 2030: The Digital University in the Next Decade*, is the outcome of 14 months of interviews, focus groups and surveys of more than 125 campus faculty, researchers and support staff. It lays out a strategic vision for UC San Diego to create the technology and staffing infrastructure necessary to support research over the next decade.

Five key themes emerged, highlighting core characteristics of research over the next decade. These themes build on current campus strengths, and are deeply informed by local, UC and national drivers.

Theme One: We will transform the process of research through innovative technologies and efficient approaches that lead to new discoveries, technologies, cures, scholarship, and creative works that improve and enrich society.

Theme Two: We will leverage our entrepreneurial spirit and build on our track record of innovation by designing, creating, and conducting research on cutting-edge technology, advancing the state of the art and partnering with industry to provide models, instrumentation and training.

Theme Three: We will generate, consume and analyze petabytes of data in our successful quest to increase scientific knowledge and improve life on the planet.

Theme Four: We will conduct research with privacy, openness, sustainability, integrity, transparency, and reproducibility at the heart of all work. We will provide an environment where diversity, equity, inclusion and accessibility enable faculty, students, and staff to excel and thrive.

Theme Five: We will leverage collaboration and interdisciplinary research on a national, regional, and local scale to apply our discoveries most effectively to the pressing problems of society.

To address the needs for campus research over the next decade, and to address the challenges posed by these themes, we make the following recommendations, detailed in the section "Recommendations: Making 2030 Possible":

1. **Make advanced computing power accessible to researchers from all disciplines**, including traditionally underserved areas such as arts, humanities and social sciences. This will allow for the synthesis and analysis of petabytes of data from multiple sources, driving new scholarship and discoveries.
2. **Make research data transfer and storage infrastructure and services available to underserved locations and disciplines** to facilitate the access and study of data.

3. **Provide advanced technology, cybersecurity, and data management training and support** to allow teams of researchers with different levels of technical expertise to work together smoothly and to protect intellectual property.
4. **Keep campus high-speed networking up to par with rapidly increasing requirements to support interactive experiences, experiments, and education for all disciplines**, including those where high-speed network has not previously been a need. Long an area of strength for UC San Diego, this will support broad participation in novel experiences and allow hundreds of petabytes of data to be accessed remotely and analyzed in real time.
5. **Fold federated authentication into workflows**, to bolster security and allow for seamless connections between instruments creating data, storage locations, and analysis tools.
6. **Provide computation, engineering, analysis, virtualization and workflow tools, environments and services** to give researchers across all disciplines, including those who have not traditionally employed computation, new ways to process, understand and share their work.
7. **Align campus research infrastructure and policy with international standards**, permitting access to local resources and immersive participation in off-campus experiences.
8. **Streamline processes and tools for researcher onboarding and research administration** to prevent impediments to collaboration between departments or divisions.
9. **Align management and compliance processes around data and research** to facilitate compliant and safe sharing and collaboration with corporate partners, between states and institutions, and across national boundaries. Clarify and streamline processes for handling/sharing PII, commercially protected data or NDA-protected information to facilitate interactions with the commercial/private sector. **Establish a robust central authority** to make it easy for faculty to get guidance on these issues.
10. **Along with San Diego City & County, build and share easily identified and accessed joint infrastructure and processes** for transferring information and guidance in both directions, with clearly designated contacts and responsible parties.

In order to achieve these recommendations, the campus will need to create a detailed implementation plan; research creative approaches to sustainably funding aspects of the plan that require it; secure buy-in from campus groups that will have to enact the plan; and regularly update the implementation plan over the decade to come.

The Road to Blueprint 2030

Introduction

UC San Diego's past investments in research cyberinfrastructure made possible enormous accomplishments, putting us at the forefront of scientific discovery, technological advancement and innovation in the arts and humanities. But neither research nor technology sits still. Evolution in the nature of research has altered the demands on IT providers. Seismic changes in the way computing infrastructure is delivered — and paid for — have collided with the way in which research is funded. Research products have broadened to include new types of "data," and regulatory and security requirements have increased, all while our commitment to the public interest has grown and strengthened.

In addition, campus leadership has taken bold steps to implement new physical resources in design, creativity and entrepreneurship. The recent opening of the Design & Innovation Building, and the parallel support for the Design Lab, demonstrate the long-term commitment to bringing new opportunities for students and researchers. A commitment to technology and services will facilitate the creation of new scholarship using these tools.

UC San Diego has long been committed to research leadership. Several key values underlie this commitment:

- We embrace knowledge-seeking and diversity of thought.
- We continually stretch our research capabilities to take advantage of technological advances that allow us to conduct research from any location, to publish scientific results that are easily reproducible, and to develop unprecedented scientific models.
- We are committed to training and developing the next generation of researchers.
- We conduct research with attention to data ethics, transparency, and research integrity.
- We contribute to the betterment of the world around us through collaboration and engagement in international, national, regional, and local activities.

With these values as drivers, we must look to the future and take intentional steps to provide infrastructure, services and staffing for the next generation of research. This document, *Blueprint 2030*, connects these values with data from many campus sources to update guidance for research technology support over the next decade. These critical investments will facilitate new discoveries and sustain the high level of achievement for which UC San Diego is known.

A rich picture of research in 2030 indicates that we must overcome some current impediments and factor in new needs as we move forward. High-speed networking — both within the campus and between our institution and others — will even more critically underpin the daily work of scholars and researchers as they engage in remote experiences, consult remote data inputs, and share models, data, knowledge, creative works and services with the world. Currently clumsy storage

solutions will require transformation with a keen eye to data security, privacy and ethics, to work with, preserve, and appropriately share data. With collaboration the norm, services and processes must facilitate working together, from the moment of researcher onboarding through data and document creation to publication of articles and datasets. The use of technologies such as high-performance computing in a multitude of novel disciplines demands training, services, and support for both people and machines. New partnerships with industry and new demands from funding agencies — and our own commitment to public service — require that we streamline the way we share our discoveries with targeted platforms and tools. Throughout all of this, we must increase support for our researchers and students so that they can continue to thrive and lead the effort to increase scientific knowledge. Creative approaches to funding these services are a must in the uncertain fiscal times ahead.

In 2020, campus administration charged the Research Computing and Data Services Governance Committee to begin the process of creating a new long-term strategic plan for research support. This document, *Blueprint 2030: The Digital University in the Next Decade*, elaborates the necessary steps we must take over the next decade to provide infrastructure, services and staffing for the next generation of research.

This report has been informed by:

1. 2020 survey of all campus researchers/faculty (see Appendix A for survey questions)
2. Interviews with key campus leaders conducted 2020-2021
3. Small faculty focus groups conducted in 2021 (see Appendix B for focus group conversation starters)

In addition, the following were also important sources:

1. 2019 researcher/faculty discussions on research support needs
2. Catalogue of UC San Diego Research Data and IT Services
3. *UC San Diego Ithaka S+R Research Study: Supporting Big Data Research*
4. AAU/APLU's 2021 report *Accelerating Public Access to Research Data*

Research in 2030 — Five Themes

Five core themes emerged from our investigations, centered on UC San Diego's core strengths and our university's strategic mission.

> Theme One: We will transform the process of research through innovative technologies and efficient approaches that lead to new discoveries, technologies, cures, scholarship, and creative works that improve and enrich society.
>
> Theme Two: We will leverage our entrepreneurial spirit and build on our track record of innovation by designing, creating, and conducting research on cutting-edge technology, advancing the state of the art and partnering with industry to provide models, instrumentation and training.
>
> Theme Three: We will generate, consume and analyze petabytes of data in our successful quest to increase knowledge and improve life on the planet.
>
> Theme Four: We will conduct research with privacy, openness, sustainability, integrity, transparency, and reproducibility at the heart of all work. We will provide an environment where diversity, equity, inclusion and accessibility enable faculty, students, and staff to excel and thrive.
>
> Theme Five: We will leverage collaboration and interdisciplinary research on a national, regional, and local scale to apply our discoveries most effectively to the pressing problems of society.

The report that follows examines each of these themes, presenting faculty views of implications for the research workflow and describing current obstacles that must be remediated to reach this future.

Theme One: We will transform the process of research through innovative technologies and efficient approaches that lead to new discoveries, technologies, cures, scholarship, and creative works that improve and enrich society.

In 2030, research processes will have a fundamentally different focus. Rather than the physical world being the direct object of study, computationally created environments will be at the center of the research enterprise. These environments will be found in many disciplines, including health sciences, the arts, oceanography, economics and fundamental sciences. They will serve not as stopgap solutions for when researchers cannot access physical environments, but as first-class objects that provide researchers opportunities to study things that aren't currently observable. These new approaches will better equip researchers to realize UC San Diego's goal of "supporting and promoting just and sustainable forms of economic development, shared prosperity, and social and cultural enrichment regionally and globally."

- Researchers will conduct research using **large, shared digital models**, constructed from decades of data collected by multiple disciplines. Researchers will be able to detect new findings that are not observable in physical space, such as new species of living organisms, new effects of climate changes, and and the impact of changes on the ocean floor to the biosphere; they will do so by leveraging powerful computation and sophisticated, collaboratively developed models across high-speed networks that make their experiences location-transparent.

- Researchers will make increased use of **shared remote instruments,** providing the ability to drive all kinds of experiments in disparate global locations using tools available in their own lab.

- Researchers will build and consult **enormous sensor networks**, with the ability to harvest and quickly analyze data from myriad sources. These networks, facilitated by development of novel network-connected devices, will provide extensive realtime data across a number of domains (e.g. temperature, precise location, current price, wildfire or seismic activity, air quality), allowing researchers to react to emergencies, to tie together previously disparate fields of study to drive new discoveries and scientific breakthroughs, and to share important data in the public interest. Dispersed throughout the campus and the world, they will also provide input to creative work in the arts and will further discoveries in the social sciences.

- Researchers and graduate students will participate in **rich remote experiences**. As previously noted, created environments will allow for deeper engagement with all aspects of scholarly and scientific focus. Just as importantly, reinforcing the University's goal of "cultivating a diverse and inclusive university community," these environments will be available to the wider academic

community. No longer will direct experimentation be accessible only to the select few who are able to spend significant time in the field or have access to highly specialized equipment. These remote experiences will touch all aspects of the academic environment, from exploration of the earth and the universe, sensors and remote instrumentation, remote human subjects research and research-related instruction.

> "Cloud-based supercomputing will become ubiquitous. Conferences will be online via Zoom or similar. Rapid data transfer will become a universal expectation."
>
> – Faculty member, Jacobs School of Engineering

Significant evolution of the current infrastructure environment is needed to adequately support this work. Changes that would facilitate greater achievement, especially in non-engineering fields, include:

1. Increased assistance for researchers and students in accessing and using computational resources that could transform their work.
2. Options to obtain support outside the research lab for servers, software, and network tools.
3. Assistance with accessing and using remote data.
4. More efficient, cost-effective options to avoid recreating technology at the research lab level. Granting agencies have historically provided only one-time funding for technology that may be needed for several decades, but both public and private cloud services are typically paid for with recurring charges. This leads to researchers building their own computing infrastructure and staffing it as best they can, often with graduate student labor rather than computing professionals, which in turn leads to security problems, technical debt, and time spent away from the research focus. We must develop creative ways of handling one-time funding for ongoing services, coupled with active engagement with funding agencies on this universal problem.

Theme Two: We will leverage our entrepreneurial spirit and build on our track record of innovation by designing, creating, and conducting research on cutting-edge technology, advancing the state of the art and partnering with industry to provide models, instrumentation and training.

Along with working with large shared models in the outside world, UC San Diego researchers in 2030 will continue to design and create technology, models, training and other experiences that are sought after by researchers and students outside our walls. UC San Diego's goal of a "collaborative and interdisciplinary research culture that advances the frontiers of knowledge, shapes new fields, and disseminates discoveries that transform lives" will be enhanced by designing, building, crafting and developing in the context of scholarship and funded research.

- We will retain a strong emphasis on **cutting-edge, local research on technology**, advancing the state of the art. This will be true both in fields such as high-performance computing and wireless technology, where our researchers have historically had national impact, and in advancing areas such as advanced manufacturing, quantum computing, extended reality, and instrumenting the natural environment and building infrastructure. Our **focus on design,** informed by the humanities and social sciences, will fold a human-centered vision into technology and interfaces. Cross-disciplinary collaborations between the arts, engineering, and fields such as health and marine sciences will lead to advances not only in instrument design but in portable technologies for personal healthcare and environmental interrogation.

- UC San Diego researchers in 2030 will create and contribute to important **scientific models that are used for simulation and experimentation**, some of which are homed at UC San Diego, taking advantage of specialized or cutting-edge hardware and software and high-bandwidth connectivity.

- From our ships at sea to our surgeons to our robotics labs, UC San Diego researchers will host appealing **interactive training, scholarship and research experiences** that draw connections from all over the world. Just as our best-in-class coursework attracts international students, these research opportunities will draw faculty, post-doctoral researchers and graduate student researchers from across the globe to gain access to unique remote opportunities.

> "Maker spaces, CNC milling and other machining processes. Electronics, controllers, 3D printed cake decoration — you name it."
>
> – Faculty member, Division of Arts and Humanities

Building on current strengths in the network and computing environments, while lessening impediments to sharing and remote access, will support this work. This will include:

1. While UC San Diego has a long history of technology innovation, advances in fields like AI, nano-technology and robotics are rapidly changing the research and funding landscape. New opportunities will require campus to be more agile in its responses to new opportunities.

2. The continued focus on creating technology locally means a need for on-premise infrastructure for that technology, including colocation space for servers, maker spaces, robotics labs and the like.

3. Close the widening gap between technology creators and users on campus to strengthen cross-campus collaborations. Work towards a shared understanding of current services and opportunities across campus departments and divisions.

4. Assist campus users in taking advantage of high-bandwidth networks on campus and beyond, especially important as significant parts of the scientific enterprise move to virtual and shared environments.

5. Rationalize campus technology and policy requirements to reduce the friction associated with accessing campus resources from off-campus locations.

Theme Three: We will generate, consume and analyze petabytes of data in our successful quest to increase knowledge and improve life on the planet.

In 2030, UC San Diego researchers will create and use data at a scale not known before. This is the result of advances in tools that create data, huge sensor networks both locally and around the world, the growth in the use of machine learning across a multitude of disciplines, and the ability to digitize information (such as large written texts, radiology images, and physical anthropological artifacts) and novel research objects in fields where such was not previously possible. These activities, in turn, will allow us to make research progress by leaps and bounds, produce findings that can be readily validated and replicated by others, train digital models that can themselves enable positive change, and support international efforts to combat climate change, understand and alleviate poverty, and develop solutions to critical health issues.

- Campus will abound with **data-producing instrumentation** such as genome sequencers, microarray scanners and miniaturized environmental sensors. Some such instruments will underpin the growing set of **core analysis services** available to a broad array of faculty, some live in individual labs, and others will be scattered across campus — or around the world, including under the ocean. Incoming data will sit in safe interim storage while it is collected and processed. Researchers and students will be able to set up automated workflows to run, tag, and process recurring analyses/data collection events. Data will automatically be preserved with an appropriate level of security and redundancy.

- It will be straightforward to **connect even multi-petabyte datasets** — whether located on or off campus — **to appropriate computing power** for analysis and integration with local research operations. Researchers will more quickly be able to turn data into discovery and application. This process will be facilitated by continued investment in high-performance computing, one of our areas of strength.

- Campus will maintain a **robust storage infrastructure** that allows for a variety of research uses, from basic lab storage of active data to multi-petabyte datasets for national and international access. This infrastructure must be designed to address concerns about data integrity, reliability, accessibility, security, and research continuity. Within this infrastructure, the campus will provide long-term preservation services to meet funder requirements as well as to ensure the intellectual capital of the university is available for future generations of scholars and researchers.

- Well versed in **data curation and metadata tagging and translation**, researchers and students will conduct analyses and additional research using local and public, private and commercial datasets. They will easily request and receive access to protected datasets.

- UC San Diego will not only be a creator of massive data; it will also become a vibrant **center for replication of scientific research**, and will be highly trusted for its own **research reproducibility**. It will be easy for researchers to locate datasets of interest in our research data library, understand how to use them because of thorough metadata, and copy them to local storage for computation and analysis.

> "[There's an] increased emphasis on data, model and code sharing, model and assay validation, reproducibility and transparency, and requirements to document these."
>
> – Faculty member, School of Medicine

> "We are looking at a data explosion in my field; growth by x30 in the next eight years."
>
> – Faculty member, Physics

Robust data services and infrastructure options that are consistently supported will facilitate this critical work. Enabling this will include:

1. Clarifying the current research storage landscape so researchers can navigate the options available to them, and quickly find the best choices for their given needs. Find ways to help researchers determine how to handle, store, preserve and share datasets, including whether local, private, or public cloud is appropriate. No-cost and low-cost storage options are needed to facilitate access to basic services by researchers in non-traditionally grant funded areas such as the humanities, social sciences and the arts as work in their fields requires more and more digital storage.

2. Storage solutions for interim or working data, along with funding solutions that allow data to persist beyond grant boundaries and including restricted as well as open data.

3. Clear and easily accessed solutions for controlled data that can be widely adopted.

4. Tools and support for large-scale data transfer to and from off-campus resources. (Tools provided within different units on campus are typically not available or scalable to all of

campus. Further, they are funded with specific units' budgets, and lack a guarantee of persistence.)

5. A stable service for research data preservation will improve continuity of research, encourage re-use of data, and increase the University's standing as researchers outside our campus replicate and build on what has been accomplished here. Funding agencies frequently require preservation and sharing of results, without guidance or funding for this effort. The Library has taken the lead in providing a service for campus data preservation, but the current infrastructure is insufficient to accommodate most campus research datasets. A coherent, well curated approach will allow us to move confidently towards the future, including adequate local infrastructure and staffing for data that is best maintained locally, partnership with funding agencies regarding national and discipline-specific data stores, and clear guidance for researchers.

6. Continued maintenance and scaling of Chronopolis, the obsolescence-resistant long-term digital archive managed by the Library for campus researchers.

7. Clear guidance, policy, and authentication/authorization approaches around accessing and using remote data. Recommended approaches to handling intra- and multi-institutional differences in compliance, privacy and security policies and procedures.

8. Continued scaling and enhanced profile for campus data curation and reproducibility resources to meet demand.

Theme Four: We will conduct research with privacy, openness, sustainability, and integrity at the heart of all work. We provide an environment where diversity, equity, inclusion and accessibility enable faculty, students, and staff to excel and provides an opportunity for all to succeed.

Research success requires wide participation, and our values demand accessibility, inclusion and attention to privacy and compliance. In 2030, the majority of disciplines — both those who engage in these activities today as well as new participants in the arts, humanities and social sciences — will employ and share large digital data sets, use machine learning, and conduct quantitative analyses. In addition, internal data from a research university of our size, handled appropriately and respectfully, will lead to important advances in science.

- With a campus legal team that stays up-to-date on evolving forms of **data and privacy regulations and usage**, researchers will be able to use and share data knowing they are following all applicable regulations. **Open data and open research** will be facilitated.
- To create an inclusive research landscape that includes all disciplines and all levels of skill, a range of **research-focused technology training** will be easily accessible by students, staff and faculty and **computation, visualization and analysis services** will be available for engagement by research projects that need these services. **Research facilitation services** will serve as the glue connecting researchers to the tools and services they need to meet their unique needs and goals.
- Our aforementioned focus on remote experiences — both from within and without — will facilitate accessibility to all. It will also help student researchers reach UC San Diego's goal of an inclusive educational experience "that develops students who are capable of solving problems, leading, and innovating in a diverse and interconnected world." Data from these experiences and others will feed into our **rich portfolio of internal data**, allowing us to help students with issues such as mental health, address the needs of faculty, staff and students with better tuned services, and create an environment where both staff and academics can flourish.

> "Students need to be introduced to how to handle data but also helped to handle it, to set things up to be reproducible from the get-go."
>
> – Faculty member, Chemistry

> "[a] matchmaking system for interested developers who have interest, enthusiasm for field and code development talent… [it] would really lift all boats if we could do this across campus."
>
> – Faculty member, School of Medicine

To better achieve these goals, staffing and training support should be focused on these needs:

1. Consistent, robust, research-focused infrastructure and services will assist researchers in meeting growing requirements around compliance, security and privacy.

2. Training researchers, research staff and students on modern technology skills required by research will allow our researchers to achieve more in a variety of disciplines, especially non-traditionally technology-focused fields. Focus on The Carpentries' offerings has been a good starting point, but more is needed, including meeting graduate students' needs to prepare them for research activities and meeting researcher/research staff needs for just-in-time training for new technology skills.

3. Supported at a base level by a 2016 commitment, Research IT Services provides researcher-facing facilitation services, and must be grown to meet increasing demand. Embedded, discipline-specific facilitation (which can be jointly funded) is the appropriate next step to ensure researchers have targeted support for new technology needs.

4. Services for computation and development and for statistical analysis and visualization are growing needs in many disciplines. These types of fee-based services allow cost-efficient support of the commitments of grants and contracts across a broad variety of fields.

Theme Five: We will contribute to the betterment of the world by leveraging the power of collaborative and interdisciplinary activities. We will focus on engagement in national, regional, and local research activities, in order to apply our discoveries effectively to the most pressing problems of society.

In 2030, research processes will be fundamentally collaborative. Shared discovery, development, invention and education will be the norm. This trend was underway in 2021, but major changes in infrastructure, funder expectations and societal demands will by 2030 make it the default choice for research at UC San Diego. Research boundaries will no longer mirror the physical boundaries of campus, with significant partnerships between public and private institutions, and broader integration with regional centers. For instance:

- Campus research will be **multi-disciplinary and multi-institutional by default**. UC San Diego's experience fighting COVID-19, where local researchers joined with researchers around the country on dozens of funded projects, took us to the next level in collaborative research and allowed us to compete successfully for large grants, pushing the prevalence of such partnerships.
- There will be a large percentage of **public/private partnerships for research**, requiring different methods of interaction at the technical, administrative and policy level. Though such partnerships have existed in the past, 2030 will see more alliances with startups and smaller companies, and we will build them into research projects from the start as intentional partners. New policies and technology tools will facilitate the often-complex issue of data sharing between the university and other legal entities.
- Researchers will be closely **integrated with the City and County of San Diego** on research/applications of research such as Smart City San Diego. Engagement with community members and citizen scientists, and building a research bridge between town and gown, will lead to broad improvements in daily life for our city and region.

> "Collaboration is key, so it's important to make it easy for our collaborators to be able to access or databases, clusters, etc., easily. Open source/data is also valued."
>
> – Faculty member

Our campus will increase researcher ability to address societal and community needs by:

1. Smoothing policy and procedural roadblocks to conducting collaborative research, including the tools through which researchers are onboarded in different departments and divisions, tools for movement of contract and grant funds between campus departments, and tools for management of contracts and grants when jointly funded with other institutions.
2. Industrial collaborations are a significant hallmark of UC San Diego, and the evolving corporate landscape will create new challenges and opportunities. Many important innovations now come from small start-ups, rather than established tech companies. Process and policies should be honed to anticipate this.
3. Sharing data and other materials with collaborators both inside and outside the university will be facilitated by addressing identity and access management issues. This will require participation in national conversations on the topic while ensuring that UC San Diego technology direction is aligned with national standards.
4. A clear and approved process (and, ideally, a single office) for handling corporate data, especially that which is shared under an NDA or with varying levels of compliance requirements, should be established/designated. This will avoid individually crafted solutions (which increases campus cost and risk).

Recommendations: Making 2030 Possible

The changes described in the above scenarios will be made possible by significant improvements in the technical, staffing and process landscape. The progress described below can only be enabled by **prioritizing upcoming research needs** in campus and unit planning and by **employing clear and practical funding and, where necessary, recharging approaches** that take into account the way in which research is funded and the critical need for data and code to persist outside the temporal boundaries of funded projects (we detail some approaches below). While discipline-specific solutions must remain the the realms of experts in their fields, **we must move away from expecting individual organizations and programs to spin up or re-invent what should be core research services, and instead dedicate stable, consistent resources for central solutions.**

In order to achieve the recommendations below, the campus will need to create a detailed implementation plan; research creative approaches to sustainably funding aspects of the plan that require it; secure buy-in from campus groups that will have to enact the plan; and regularly update the implementation plan over the decade to come.

The following steps will help UC San Diego realize this vibrant future:

1. **Expand access to advanced computing power to researchers from underserved disciplines**. This will allow for the synthesis of massive data from multiple sources, driving new scholarship and discoveries, and allow faculty to focus on their research rather than building infrastructure.
    a. Provide basic, entry-level computational resources open to all researchers (students, faculty and staff) at no cost. This is necessary to support non-traditionally grant-funded areas such as the arts, humanities and some social sciences, and to encourage uptake of new research approaches—often a springboard for more advanced research and increased grant funding.
    b. Make available and accessible appropriate “next step” high-performance and high-throughput computing offerings at reasonable cost to those with greater computational needs.
    c. Sustain and advance support for on-premise shared compute and storage space for campus researchers and technology creators to meet tech development and privacy/compliance needs.
    d. Recruit professional system administrators and developers with an understanding of the research process, to meet the intermittent needs of individual research groups. Develop a financial model to sustain these hires and intermittent needs.
2. **Provide research data transfer and storage infrastructure and services to underserved locations and disciplines** to facilitate the access and study of data.

    a. Identify and ensure availability of centrally managed stable, secure, and compliant storage capable of handling a variety of data types, including big data, controlled data, and high throughput.
    b. Maintain data transfer nodes in major campus network nodes to facilitate the download, transit and management of large datasets.
    c. Select and commit to standardized tools for moving research data safely and securely.
    d. Support and augment a robust research data curation and stewardship program, including both staff and funded preservation and persistent (between-grant) storage solutions.
3. **Provide advanced technology, cybersecurity and data management training and support** to allow teams of researchers with different levels of technical expertise to work together smoothly and to protect intellectual property.
    a. Ensure instructional staff are available to researchers across campus, and provide consistent programmatic support for relevant topics.
    b. Develop expanded training programs to advance the technology skill sets of our student, faculty and staff researchers.
    c. Develop just-in-time training opportunities aimed at researchers and lab staff for skills required by funding agencies such data management, reproducible science, and compliance.
    d. Continue expansion of UC San Diego's Research IT team's facilitation program, including establishing more co-located support staff in departments and divisions, will provide discipline-tuned support with centralized efficiency.
4. **Keep high-speed networking at par with rapidly increasing requirements to support interactive experiences, experiments, and education for all disciplines**, including those where high-speed network has not previously been a need. Long an area of strength for UC San Diego, this will support broad participation in novel experiences and allow hundreds of petabytes of data to be accessed remotely and analyzed in real time. Collaboration will be enabled by infrastructure that allows very large datasets to be shared both within campus and between institutions in a fast, efficient way.
    a. Continually increase the intracampus network speed, along with upgraded bandwidth to buildings, with available speed increments in connections to offices and labs, keeping pace with the best-in-class speeds and technology at any given time, retaining the longstanding standard for services under the Next Generation Network (NGN) program.
    b. Continue to grow the campus' bandwidth to the national and international research Internet, to keep up with the needs of research and instruction.
    c. Upgrade wireless technology as new options reach reliable production.
5. **Fold federated authentication into workflows**, to bolster security and allow for seamless connections between instruments creating data, storage locations and analysis tools.
    a. Incorporate InCommon, a national standard for federated authentication, into all services that may be accessed by external individuals.

    b. Complete efforts towards a better-designed identity & access management (IAM) system, and fold other national standards for identity, such as ORCID, into records for UC San Diego individuals.
6. **Provide computation, engineering, analysis, virtualization and workflow tools, environments and services** to give researchers across all disciplines, including those who have not traditionally employed computation, new ways to process, understand and share their work.
    a. License critical multidisciplinary research tools, make them easily available, and support their use, with robust training to allow researchers from many disciplines to engage in advanced data analysis.
    b. Support maker spaces, robotics studios, and related resources, available for reservation and use by both researchers and students.
    c. In response to researcher demand, stand up appropriate for-hire core services to meet computation, analysis, virtualization and workflow needs.

These recommendations are not presented in order of priority. Indeed, the recommendations interact with each other in fundamental ways — infrastructure and tools alone solve nothing. They require adequate staffing support to align them with research needs and to integrate them into the research workflow.

Similarly, alongside this technical growth must come reductions in workflow friction:

1. **Align campus research infrastructure and policy with international standards**, permitting access to local resources and immersive participation in off-campus experiences.
2. **Streamline processes and tools for researcher onboarding and research administration** to prevent impediments to collaboration between departments or divisions.
3. **Align management and compliance processes around data and research** to facilitate compliant and safe sharing and collaboration with corporate partners, between states and institutions, and across national boundaries. Clarify and streamline processes for handling/sharing PII, commercially protected data or NDA-protected information to facilitate interactions with the commercial/private sector. **Establish a robust central authority** to make it easy for faculty to get guidance on these issues.
4. **Along with San Diego City & County, build and share easily identified and accessed joint infrastructure and processes** for transferring information and guidance in both directions, with clearly designated contacts and responsible parties.

Funding of these growing needs will require innovation and flexibility. Although an oversight team (discussed below) must guide this ten-year plan through realization and will need to determine the appropriate sustainability steps along the way, some approaches include:

1. **Grant funding**, a traditional resource for building infrastructure, is helpful for breaking ground on large, shared computing and network infrastructure. At UC San Diego, grant

funding has brought us large supercomputers and cutting-edge campus Internet gateway speeds. Science drivers abound for these efforts, which are most useful when the infrastructure is novel or cutting edge in nature, is expected to last for at least five years, and serves a broad swath of campus faculty and the surrounding community.

2. **Negotiation and discussion with granting agencies** should be ongoing. Grantmaking habits combined with the move towards cloud have left many campuses struggling to maintain research data over the long term with one-time monies. The agencies have an interest in maintenance and sharing of research outputs, and as a large R1 we have a voice at the table that should be employed.
3. Technology loans, gifts and try-outs from **external companies** can allow researchers to experiment with novel technologies, pushing research forward at minimal cost to campus. These efforts work best when central IT, divisional or departmental IT, and researchers work together to integrate the technology with existing campus infrastructure and can look to ways to incorporate the new facilities in the longer term.
4. **Pilot offerings** from technology providers can supply excellent discounts in early years of use, but require committed faculty willing to try out the technology within a relatively short window, or risk lying unused.
5. **Partnering with companies for development** gives campus a voice into the shape of commercial technology. These opportunities require a clear commitment of campus resources (both in the lab and in central IT) to realize.
6. **Commodity services** that are available in the cloud or as licenses may be made available for purchase by individual faculty who require them through special licensing or discount arrangements negotiated by the campus as a whole.
7. Related, **reprioritizing** may be an important part of arranging support. As we look towards the future, we see a need for a lowered emphasis on maintaining local commodity equipment and services locally, and a higher emphasis on applying university resources to research-specific, discipline-specific, hands-on, experimental, cutting-edge, and creative technologies.
8. Finally, **core funding** should be directed where it is most critical — funding services that either (1) meet the needs of non-traditionally-grant-funded disciplines, (2) meet universal basic needs of research, and/or (3) protect core interests of the University itself (e.g. intellectual property).

Appendices

Appendix A: Researcher / faculty survey questions:

1. What trends are you seeing from your key funding agencies? (For example: changes in frequency and deadlines for funding; open data requirements; a focus on collaboration, or on multidisciplinary research; etc.) In what ways might they push us to provide research data and computing services differently than we do today?

2. We're trying to plan for the next five to ten years of support for data and computing services. What do you see as possible developments, needs, or changes in research processes/workflow in your field over the next few years that you'd like us to consider?

3. What are you doing manually or laboriously today that you would like technology to help you with?

4. What is the campus doing well to support research computation and data management?

5. What inhibitors are currently in place at UC San Diego in the research computation and data management space?

6. In what ways are you collaborating with colleagues at other institutions? What about collaboration with industry?

7. What are your colleagues at other institutions doing that you're not able to do here?

8. What keeps you awake at night, either because it scares you or excites you?

Appendix B: Focus group discussion starters:

Focus Group One: Collaborative research - impacts, needs and issues

Contemporary research relies on collaboration at many levels, from inter-lab work on campus to large international efforts. Further, UC San Diego has made research collaboration a priority for many years. However, there are still many roadblocks and complications to effective and efficient collaboration. This focus group will look at how campus can improve support for collaborations at all levels.

FOCUS GROUP KICKOFF QUESTIONS:

1. In the next 5-10 years, where do you see funding agencies focus and connected demands?
2. What challenges have you faced in the past in the area of collaboration (inter-lab, international, etc)?
3. What would you like to see campus do to support multidisciplinary and/or multi-institutional work in terms of infrastructure in the next 5-10 years?
    a. (maybe) Would these things be helpful for intradepartment or intra-lab work as well?
4. What technologies or approaches have you been able to adopt or adapt to meet your collaborative needs?
5. What are the origins of any collaborative challenges when working with industry - is it coming from the campus side or the industry side?

Focus Group Two: The role of data in future research processes

Data is the driving intellectual capital of the contemporary research university. Support and infrastructure for data creation/collection, storage, sharing and preservation has become complicated and expensive. Policies and requirements around data have proliferated, from local and UC policies to funder requirements and domain expectations. This group will look at how the campus should move into the next generation of support for research data.

FOCUS GROUP KICKOFF QUESTIONS:

1. What do you expect your data problems to be in the next 5-10 yrs? w.r.t.
    a. Volume/expected growth
    b. Collection & transfer (ingest)
    c. Access & processing
    d. Archiving, preservation & cataloging
    e. Sharing (incl. publication)
    f. Compliance
    g. Security/disaster recovery
    h. Training & support
    i. Anything we forgot?

i. What is currently a challenge for you? How are you solving these problems today? What's still unsolved?

Focus Group Three: Emerging/increasing research foci

The contemporary university has seen a rapid increase in all aspects of the research process, from tools and infrastructure based on radical new technologies to a wide variety of outputs, with expectations of extensive intelligent computational analysis. This group will look at how campus can provide appropriate support at all levels for potentially unknowable next developments. [maybe mention some specifics like machine learning, code & data as outputs, convergence of research & instruction issues]

FOCUS GROUP KICKOFF QUESTIONS:

1. In the next 5-10 years, where do you see funding agencies focus and connected demands?
    a. How are funding agency priorities on specific topics or processes (such as data science, machine learning, AI) changing your research?
2. What do you expect your needs to be around access and preparation of data to be used within ML contexts? Includes interfaces, security and access.
3. In the next 5-10 years, with the changes in methodology and increase in data, what will your most important research outputs be?
4. What kinds of training, services and support will be needed to accommodate the increased relevance of data science and machine learning in nontraditional disciplines?
5. In what ways will trends for instruction drive some of the research trends, ways that we engage with students? What about the reverse? (This isn't just a funding agency issue, but also demands from our campus.)
6. What would make it easier to get started with new technologies/infrastructure that you may need to support your research?

Appendix C: Blueprint for the Digital University (2009):

Blueprint for the Digital University