Case Study

Judith A. Pirani, EDUCAUSE Consultant and President, Sheep Pond Associates

Highlights

Related ECAR Research Report: Research Computing: The Enabling Role of Information Technology

Case Study Institution: University of Washington

Issue: Practices that facilitate the development of supporting technology solutions for researchers

Solution: The University of Washington and the eScience Institute have developed a robust portfolio of tools and services to support research activities. Over time, several practices have emerged that any institution can utilize in similar research support efforts.

Contact: Cara Giacomini (cgiacomini@uw.edu), Research Manager, Academic & Collaborative Applications, UW Information Technology

The University of Washington (UW) is a leading research institution, receiving more federal research funding than any other American public university, garnering more than $1.5 billion in total grants and contracts in 2011. The UW Information Technology (UW-IT) department and UW’s eScience Institute (an externally funded organization that assists researchers in applying technology for data-driven research) have been charged with supporting the research community’s vast and sophisticated technology needs. Both organizations have responded accordingly, incorporating technologically innovative solutions in areas like data analysis, high-performance computing, cloud services, digital asset management, collaboration, and high-speed data networking to facilitate researchers’ activities.

Best Practices

- Conduct ongoing research and outreach to understand user needs
- Develop tools and services for use in multiple environments (e.g., teaching as well as research)
- Balance central and local support
- Establish funding with sustainability in mind
- Pursue partnerships and collaborations

Benefits

- Centrally supported research tools and infrastructure enabling researchers to focus on their research rather than IT maintenance
- More efficient development and maintenance of resources
- Development based on documented needs instead of conjecture
Case Study Inspiration

The University of Washington’s research support practices mirror recommendations presented in ECAR’s study *Research Computing: The Enabling Role of Information Technology*, most notably that “IT should be proactive in uncovering research computing needs and communicating which research services are available.” The case study uses UW’s research support to illustrate this recommendation.

Although many institutions lack the resources—or even the need—to offer similar levels of support, at the heart of UW’s activities lie several practices that any college or university can apply when supporting its own research communities, regardless of the size and intensity of institutional research activities. This case study outlines UW’s practices and how the UW-IT and eScience organizations apply them to solve researchers’ problems.

**Background**

Today’s research technology environment is a world of huge data sets, virtual collaboration environments, and a growing repertoire of analytical tools. Its scope expands beyond the physicist or astronomer working with huge data sets and simulations to the humanities professor who stores, catalogues, and tags images or sounds. Data sources continue to multiply: online surveys, Facebook data, scanned 19th-century maps, and digitally recorded bird songs are some examples of this variety of data. A researcher is just as likely to collaborate with a colleague down the hall or across the globe, and cross-disciplinary research is also increasingly common. Consequently, these changes and the complexity they add have led more researchers to discover that their traditional research tools—desktop computers, local servers, spreadsheets, individual databases, and statistical software—aren’t scaling to today’s research needs.

In 2007, UW-IT and the eScience Institute staff recognized this change and set out to discover how to best support their researchers. They conducted a series of interviews with a cross-section of top UW researchers, regardless of field and experience. They found that researchers were collecting significantly more data but storing it on their own computers, thumb drives, or back-up drives, and they analyzed data using approaches that had not kept pace with the growth in volume, rate, and dimensionality. They needed more exposure to data management best practices, as well as tools and services to assist their research activities and collaboration with others.

**Practices**

Since 2007, UW-IT and the eScience Institute have used their findings to become trusted, knowledgeable partners by empowering the researchers through technology, providing relevant tools, services, and support. Over time, both organizations broadened support beyond the sciences to humanities and social sciences, and they deepened support across all fields. Interviews with UW-IT and eScience Institute staff revealed five facilitating practices that have emerged from their ongoing activities—needs awareness, understanding of their

---

*Based on [our interviews], we adopted this philosophy that data-driven discovery would be pervasive across all fields, and we had to somehow figure out how to distribute the capabilities widely across campus or we would stop being competitive.*

—Ed Lazowska, Founding Director, eScience Institute, on initial research
user base, development of multipurpose tools, balancing central and department support, and sustainable funding.

**Needs Awareness: Outreach and Community Building**

The 2007 research set in motion an ongoing outreach effort to the UW research community. Upon publication of UW-IT and eScience Institute staff members’ ECAR research bulletin, “Information Technologies for eScience: A Preliminary Report from the University of Washington,” UW-IT held a campus-wide “road show” to present the findings and to engage researchers about data management problems. Today, UW-IT conducts focus groups and surveys to gather formal feedback. Staff members regularly interact and listen to faculty, attending meetings of various groups and participating on various e-lists to explore unmet needs and understand how research fits into other faculty roles. As UW-IT focuses more closely on the humanities and social science researchers, it employs its original 2007 research methodology, probing faculty for the type of data collected, how data are stored, ongoing challenges, and support requirements.

The eScience Institute reaches out by re-creating the computing center’s social hub. Historically, computing centers not only housed facilities but also provided a valuable social forum, like the local coffee shop or water cooler, where people could share ideas and help each other, especially novices. In the past, the migration of computing from centralized facilities to individual desktops or, more recently, third-party commercial service providers has diminished the computing center’s role. Technologically, this is not necessarily a bad thing, but that important social element was lost in the transition away from the use of centralized technologies. The eScience Institute has made an effort to reestablish this focal point, forming a cadre of scientists in a range of fields to help researchers write and/or execute grants that utilize data-driven discovery techniques in their research. One way the institute fosters this community is to host seminars, workshops, and brown-bag lunches. Another is through its website. Researchers can e-mail the institute with a problem; other requests come from personal interactions. Regardless of the source, eScience staff members triage the next steps for every request; it could be an advisory conversation, technical contributions to software development or algorithm design, or a longer-term research collaboration that tackles a broader problem.

Both UW-IT and eScience Institute staff emphasize the long-term nature of their outreach activities. Acceptance doesn’t occur overnight. Both organizations spent years building community awareness, but today both organizations report that more and more researchers are eager to reach out to them with new ideas or requirements.

**User-Based Assessment: Understanding the Research Pyramid**

Several UW-IT and eScience Institute staff described UW’s research environment as a pyramid (see Figure 1). At the top are the megaprojects, which receive significant funding

---

We are not walled off in an ivory tower. The campus computing community is a great resource for us. We meet them constantly on many different fronts, from governance to budgeting to tools. I view it as a very collaborative ecosystem. Research support is not all necessarily focused and driven by central IT.

—Brad Greer, Assistant Vice President, Computing Infrastructure, UW-IT, on outreach and community building
and sustain their own research technology needs. The research population broadens as one moves down the pyramid, and technology adequacy may lessen too. Researchers may inherently recognize the need to collect, manage, and analyze their data in new ways but don’t know how to proceed. This can be especially true for humanities or social sciences faculty who may not be exposed to as many different technologies and thus may be unaware of their options.

A UW-IT staff member described the support implications. He estimates that a decade ago, about 15 researchers required access to scalable high-speed computing resources on campus, and in the coming decade, this number could balloon to more than 1,500 people. In addition, there are the one or two thousand users who lie outside the pyramid, the faculty members or graduate students who may not recognize the need for research support or do not even know to get started. This doesn’t translate into a need for 100 times more support resources at 100 times the cost, but it does illustrate the necessity of ramping up research support resources to handle this increasing demand.

We characterized it as moving from lab notebooks to spreadsheets to databases to cloud data mining and visualization. One of the biggest hurdles for the novice user is to understand how something could apply to their problem.

—Erik Lundberg, Assistant Vice President, UW-IT, on understanding the research pyramid

![Figure 1. The Research Pyramid](image)

**Multipurpose Resources: One Tool, Many Uses**

When UW-IT and the eScience Institute talked with the faculty members, it became obvious that teaching overlapped their research activities. But typically teaching and learning tools were in one walled garden and their research collaboration tools in another.
Now UW-IT and the eScience Institute try to develop tools that bridge multiple domains and disciplines. UW-IT’s involvement in multiple areas (e.g., research, business computing, student computer use, and teaching and learning) enables them to pinpoint technologies or skill sets that can help broad constituencies. The eScience Institute’s interdisciplinary approach offers similar advantages. Here are several examples:

- The eScience Institute developed a cloud-hosted tool for collaborative access to research data; this tool, SQLShare, is now being used in classrooms to deliver research data to students for assignments and projects.

- An architecture department member suggested a mobile app to photograph things in the field and then to annotate those pictures, display them in a collection, and share them. Although the tool was envisioned initially for architecture, UW-IT discerned its broader applicability and wants to build the mobile app for general use.

- UW-IT recently implemented the Tegrity lecture capture tool. Though designed for teaching and learning, the tool has broader potential (e.g., research team information sharing, faculty workshops, and guest lectures). Because access is tied to an official class of record and a specific user space, UW-IT built a tool that creates spaces in an artificial course to allow non-class-event capture.

- With cloud-based applications like the new Canvas learning management system (LMS), Office 365, and Google Apps for Education, UW-IT implements and integrates them so that the tools support collaboration and communication in teaching and research contexts.

Central vs. Department Support: Balancing the Benefits of Both

UW-IT and the eScience Institute recognize that central and department IT organizations should work hand-in-hand to support researchers’ technology needs. At UW, most IT activities occur in the departments, but centralizing infrastructure, services, or standard platforms, such as an LMS or messaging services, offers more efficiency when scaled for campus-wide use. So whereas UW-IT and the eScience Institute serve as focal points for broadly available resources, local IT organizations provide domain-specific knowledge (e.g., how to apply a technology to meet a particular goal or how to tackle intricate, discipline-specific data management issues). Providing a solid, centralized base of research support tools and services frees up researchers to focus on their science and frees up local personnel for domain-
specific support. Maintaining that link between central management and local departments is vital so that each is in tune with the other’s activities. UW-IT and the eScience Institute’s outreach activities strengthen this connection.

**Funding: Self-Sustaining Strategies**

The economic crisis impacted UW’s state funding allocation, which in turn curtailed available resources for research support. Consequently, UW-IT charges a service fee to sustain some internally built and hosted technology and services (e.g., servers and data centers). It also aggressively forms partnerships with higher education entities and with corporations to make cloud and other externally provided services more affordable—or even free—to researchers, who pay for those services through grants and/or department funding.

**Supportive Services and Tools**

UW-IT and the eScience Institute applied these five principles to develop technology tools and services to assist researchers. Following are several examples of how UW-IT and the eScience Institute identified researchers’ requirements and developed corresponding components of their research infrastructure. The notes offer links to websites that provide more information.

**Data Analysis: SQLShare**

**Issue:** Conventional databases can be problematic to operate, requiring a combination of training, adaptations to the technology itself, and handholding to help beginners.

**Solution:** The eScience Institute’s SQLShare is a cloud-based database service, eliminating the need to install and maintain database software locally. A researcher simply loads the data, without setup or schemas, and can immediately start querying and sharing the data. SQLShare acts as a “delivery vector” for a specific set of database features that are especially useful for researchers, while eliminating the features they typically don’t need. The result is a simple and flexible platform that has led to applications in environmental microbiology, chemical engineering, observational oceanography, and cancer research. SQLShare has been used on the UW campus and beyond for teaching, for publishing data associated with papers, for interuniversity data sharing and collaboration, for exploratory data analysis by nonprogrammers, and as a drop-in replacement for conventional databases.

**High-Performance Computing: Hyak**

**Issue:** A researcher may find her work requires more processing power and storage than her desktop solution can handle, but migrating to a larger-scale computing system requires significant investment of equipment, time, and resources.

**Solution:** An alternative is UW-IT’s and the eScience Institute’s co-developed Hyak, a centrally managed, shared, high-performance computer cluster. Hyak is financed through a condominium business model: Departments invest in a certain amount of capacity in

---

In practice you get about twice as much work done per dollar that you spend in Hyak versus building it yourself. There are obviously technical challenges to building Hyak. But just as important is the creation of a rational business model with sufficiently low prices to incentivize researchers to participate, that doesn’t bankrupt the central administration, and where everyone feels that it was fair.

—Chance Reschke, Research Consultant, UW-IT, regarding Hyak
Hyak, into which the researchers purchase and install their CPUs. The deans and senior administrators invest in a durable system infrastructure, and the researcher pays a low rate for their CPUs and memory, eliminating the need to install and maintain their own systems. In addition, the shared resource offers more than 100% of a researcher’s purchased computational capacity—Hyak executes each researcher’s work using his own CPUs and any available spare cycles.

Cloud Services: NET+

Issue: Cloud-based infrastructure, software, and communications services reduce researchers’ local IT support and maintenance requirements, but it can be a time-consuming as well as costly proposition for an individual school to evaluate and select appropriate services.

Solution: To offer researchers a less expensive and easier way to adopt cloud services, UW is a participant in Internet2’s NET+ cloud service initiative. Together, members leverage their collective bargaining power to align higher education needs with vendor offerings.

Digital Asset Management and Collaboration: MediaAMP

Issue: Researchers grapple with how to store and manage their digital assets (e.g., videos, digital photographs, recordings). In addition, more researchers want to repurpose their digital assets for collaborative opportunities, as well as for research and educational purposes.

Solution: UW-IT created MediaAMP, a hybrid system offering a comprehensive, flexible, and extensible service for data asset management, content delivery, and global storage. To create this solution, UW-IT forged public and private partnerships, bringing together best-in-class commercial frameworks and open-source cloud services to offer capabilities that especially appeal to researchers. For instance, MediaAMP offers encrypted live or archived video so researchers can share data in HIPAA-compliant ways. UW-IT can use MediaAMP in less obvious ways also, such as embedding it in UW’s Canvas LMS to manage course videos and to easily insert videos and collaborative events into course curriculum.

High-Speed Data Transmission: Science DMZ

Issue: Campus firewalls filter out Internet threats, but they also create choke points that inhibit sharing research data. In theory, UW’s network can move data between campuses at more than 1 gigabyte per second, but in practice, firewalls and desktop computer circuit size limit throughput to about 5 megabytes per second, less than 1% of the network’s potential and far too slow to accommodate large-scale data transfers in a timely manner.

The idea is to provide an immediate service [of holding a live web event], but what is always interesting to the researchers is to collect, extrapolate, or repurpose what they have done—and to have a way to share with others.

—Laurens D. Baker, Director, Digital Media Platforms, UW-IT, regarding MediaAMP

Start [your research support activities] by designing a science DMZ to reduce the friction when moving data between end points. When you invest the time and effort in designing your science DMZ correctly, everything else is easier from that point.”

—Chance Reschke, Research Consultant, UW-IT, regarding a science DMZ
Solution: A science DMZ is a high-speed network subset that accelerates throughput to specifically linked machines. It directly connects specific high-capacity research computing and storage systems and researchers’ workstations either on campus or to similar resources at other institutions, eliminating transmission bottlenecks.

Next Steps

In the future, both UW-IT and the eScience Institute will work to increase the number and scope of researchers who use their services and tools. As more researchers heed the call to ramp up their research with advanced technology, the next step is to scale up technology services to help more researchers. In addition, both organizations will continue to develop multipurpose tools that address problems in both teaching and research contexts.

Also, both organizations want to adopt more cloud services. They are talking with UW researchers directly and co-sponsoring workshops with Internet2 and Microsoft to engage broad members of specific research communities (e.g., social sciences) to understand their research and scholarship requirements. The latter initiative will launch a series of exploratory and developmental projects. Efforts continue with vendors directly or through NET+ to negotiate access and price rates to cloud services and to fine-tune those services for researchers’ specialized needs, such as HIPAA protection for clinical data research.

Educational efforts persist too. UW-IT will continue to share its research findings and general information to foster higher levels of awareness around campus about smart data practices. The eScience Institute wants to develop students’ expertise in data-driven discovery through certificate programs and graduate degree programs. The institute submitted a proposal with the NSF’s Integrative Graduate Education and Research Traineeship to create a graduate program in large-scale data analysis. One UW-IT staff member hopes collaborations like Internet2’s Cloud Services Working Group will lead to public forums to share information and best practices across all institutions, both big and small.

Lessons Learned

Conversations with UW-IT and the eScience staff highlight several lessons to pass on to other IT organizations that plan to enhance their institution’s research support.

- Reach out to the researcher community. The initial 2007 researcher interviews pointed the way for UW-IT and the eScience Institute’s technology support activities, but continued contact ensures ongoing understanding of researchers’ technology requirements and offers a conduit for researchers to suggest new tools and services. It doesn’t matter how central IT reaches out—a listening tour, brown-bag lunches, or a special interest group about research-related technology activities are all good methods. Outreach built trust among UW researchers and UW-IT/eScience Institute staff, creating a willingness to turn to central sources for support instead of tackling problems themselves.

Every institution will need to get these services up for researchers, and these forums can help develop some of those processes.

—Erik Lundberg, Assistant Vice President, UW-IT, on reaching out to the researcher community
• **Build momentum gradually.** Launching a research support program requires huge activation energy because every problem is a new problem, requiring considerable central IT resources to tackle it. But over time, experience builds a knowledge base from which central IT can draw to solve subsequent problems more efficiently, eventually freeing resources to expand support operations. In addition, a slower ramp-up checks possible overextension of resources, which is especially prudent in these unpredictable funding times. For example, UW-IT and the eScience Institute opted to roll out Hyak in stages, in response to local demand, to avoid building an underutilized system. Phase 2 is under development now, and if Hyak remains popular, they will add a third phase, rolling it out even more judiciously, on a rack-by-rack basis instead of in previous increments of 20 racks.

• **Balance today and tomorrow.** The research environment continues to change rapidly, and central IT needs to be cognizant of how technology, cloud resources, collaborations, and other institutions’ activities may affect researchers’ current needs. In addition, this environment must be open and flexible to adapt as new solutions emerge or as user requirements change. The key is to invest resources to address current needs while remaining agile enough to respond to future changes.

• **Build people skills.** Technology may gain the limelight in research support, but the underlying people skills are important, especially as research becomes more complex, involving bigger data sets, faster networks, and a wider variety of tools. Many solutions are customized and unique, and central IT will need the people skills to implement them.

• **Understand your capabilities, and look to other resources as needed.** Research technology support needs are accelerating during a time when central IT may experience limited resources. UW-IT and the eScience Institute understand their organizations’ capabilities and when it’s appropriate to turn to consortia and other resources—such as Internet2’s NET+ cloud services or the InCommon identity management federation—to provide more resources at better costs and easier institutional access instead of trying to build all their solutions themselves. Alternatively, they might plug in outside resources to enhance homegrown solutions. For example, SQLShare’s cloud hosting translates into less than one developer’s time to maintain it since its inception, and the commercial cloud provider has the resources to automatically recover from any service blip almost immediately—far faster than UW staff could fix it manually. Outside resources, in turn, free staff to focus on value-added activities.

• **Don’t neglect the business sector.** The whole push toward big data affects both research and industry and has led to common data analysis practices and tools. Institutions should evaluate solutions that originate in industry as well as the higher education community.
Stick to common standards. UW-IT and the eScience Institute’s commitment to open and transparent standards, solutions, and definitions whenever possible fosters the plug-and-play model of internally and externally developed services and tools into their research support portfolio.

As central organizations like UW-IT and the eScience Institute continue to address researchers’ solvable problems—such as setting up a database—with today’s technologies, there is less waste of efficiency, freeing up researchers’ mindshare to tackle today’s and tomorrow’s harder—and more interesting—problems.

Acknowledgments

ECAR would like to thank the following for their assistance, insights, and time in researching this case study: Laurens D. Baker, Director, Digital Media Platforms, UW-IT; Cara Giacomini, Research Manager, Academic & Collaborative Applications, UW-IT; Brad Greer, Assistant Vice President, Computing Infrastructure, UW-IT; Bill Howe, Director of Research, Scalable Data Analytics, eScience Institute; Ed Lazowska, Founding Director, eScience Institute; Tom Lewis, Director, Academic & Collaborative Applications, UW-IT; Erik Lundberg, Assistant Vice President, UW-IT; and Chance Reschke, Research Consultant, UW-IT.

About the Author

Judith A. Pirani (judith.pirani@gmail.com) is an EDUCAUSE Consultant and President of Sheep Pond Associates.

Citation for This Work


Notes

2. See http://escience.washington.edu/.


7. See http://escience.washington.edu/content/hyak-0.


10. Recently, UW-IT and the eScience Institute hosted—a workshop called “Cyber-Infrastructure for the Social Sciences.” This was one of a series of workshops to explore ways in which public cloud infrastructure can support big-data needs in research. This workshop, held on the University of Washington campus, brought together researchers from a variety of social science disciplines from a number of universities, along with CIOs, to discuss the unique challenges facing those fields, as well as possible common approaches to address those challenges. For more information see http://i2azure.com/Past_Workshops.html.