Providing Services for the Enterprise vs. Supporting Research

Ideas for discussion, by Rick Jaffe

As the Research Information Technology (RIT) group moves towards providing services to UC Berkeley's research community, what lessons can we take from our experiences in the service business? In particular, how will we balance the needs of central IT, with its broad responsibilities and competing demands for limited resources, with those of the individual faculty, post-docs, graduate students and other scholars doing research on campus? When do we aim for the “sweet spot” of general functionality, and when do we provide the focused attention and expertise to solve particular problems? How do we build the trust of our campus research partners – our “user base” – that will be necessary for a sustained run?

Research Hub is an online collaboration service that has been offered to campus by IST for nearly three years. As a generalized collaboration tool with aspirations to be of particular service to researchers, the experiences of the “hub” can provide insights and raise important questions. Basic decisions in strategy, design, and operation – alignment with enterprise requirements, scope of service, choice of platform or vendor, willingness to customize – have shaped Research Hub’s ability to deliver as a service.

In a number of ways, being an enterprise service has prevented Research Hub from fulfilling its promise to campus users. The divide between the concerns of central IT and those of campus researchers is deeply organizational, growing from differences in goals, priorities, timelines, budget practices, responsibilities and rewards. For RIT, navigating this divide may well turn out to be a much more defining endeavor than finding the most appropriate technological solutions.

Table of Contents

Research Hub background..................................................................................................................................................2
Service Strategy ......................................................................................................................................................................2
Service Design .......................................................................................................................................................................4
Case studies: User needs and how Research Hub has attempted to meet them ........................................6
Conclusions........................................................................................................................................................................9
Providing Services for the Enterprise vs. Supporting Research

Research Hub background

Research Hub is an online collaboration environment that is essentially a re-branding of the Alfresco Enterprise Content Management system. Alfresco has brought expensive ECM functionality to the open-source world so that it is affordable by smaller enterprises. The product is designed with “openness” as a value: it is standards-based and extensible, and exposes application program interfaces (APIs) that allow developers to build additional services.

Research Hub is easy to use. No arduous training is required. Content is accessible via web browser, WebDAV (allowing bulk uploading and downloading of folders and files, and syncing with a local drive), mobile app (using the CMIS standard) and IMAP (for integration with email clients). Campus provides 10GB of storage quota free to each individual; additional storage can be purchased. Groups can establish departmental accounts to maintain access to content even after the individual who uploaded or created them is no longer affiliated with campus.

Research Hub attracted some 2000 new users in each of its first two years of operation. Eight months into this year, that number has grown by 2500 or so. (The total number of people who have logged in at least once to Research Hub now stands at 12,500, but this count includes a large number who have logged in by WebDAV simply to download applications from Software Central, which began to use the hub for this purpose about nine months ago.) With the decommissioning of the AFS (Andrew File System) service and the phasing out of bSpace (the campus’s learning management system), Research Hub has seen an apparent increase in use over the past months by units within academic areas of campus.

Service Strategy

From the beginning, Research Hub’s service strategy has been closely tied to other offerings in the collaboration/file-sharing/document management field. The service came into being at a time when IST wanted to strengthen its support for collaboration, but faced the prospect of being tasked with supporting not just the services it already offered to campus but also a wave of modern options newly available on the Internet. Research Hub was a solution targeted at a broad segment of the campus, with the intent of simplifying the service portfolio.

At the same time, the Alfresco ECM product presented IST’s Research Content & Technologies (RCT) group with the opportunity to try a new approach to providing services. The repository and its interfaces would serve as the underlying platform; the “Share” application (branded “Research Hub”) would be the user-facing service on top of the repository platform. Other services, such as records management,
Providing Services for the Enterprise vs. Supporting Research

were identified as future additions to this “hub.” Tools, data sources and services commonly used by researchers could be integrated into the platform, as well.

To demonstrate how the collaboration service/platform might address the needs of campus researchers, Research Hub staff created a prototype service that connects the Research Hub with the California Digital Library’s UC3 Merritt Repository for data preservation and access. The proposed service enables “one-click” depositing of materials from the Research Hub to the Merritt repository and, in return, the assignment of a permanent URL for public access to those materials.

Yet Research Hub has never been closely identified with a concrete base of users. Emerging from the Media Vault Program, its trajectory was quite the opposite: RCT sought a “low-touch” solution for as wide an audience as possible. In the ensuing years, Google and Box have arrived in the campus-supported collaboration “space” to fill that role. These cloud services are managed by others and have the funding and revenue base to drive improvement. Adoption of these services has been swift; overshadowed and outcompeted, the Research Hub initiative has taken a back seat.

Similarly, progress on the Alfresco platform strategy has stalled. Efforts to sell the envisioned suite of services by focusing on the financial, environmental and risk control benefits of a broad records management program have not yet proven to be compelling to campus decision-makers.

The Merritt Repository integration remains Research Hub’s strongest calling card. Ready for piloting and assessment, but waiting further funding, this integration serves an early model for future development.

Lessons we might learn from the Research Hub experience include:

- **Efforts to rationalize the broader set of services can run counter to support for on-the-ground needs.** Though the Alfresco platform is open-source and customizable, therefore capable of being tailored to fit campus needs, tying the development of the Research Hub program to the larger IST and campus IT picture has obscured the value of this singular asset.

- **Moving ahead without strong campus sponsorship or serious stakeholders within the campus community can be a mistake.** Offering a service to campus members – eliciting their trust in order that they expend the resources of time and effort to change to a new set of tools – without truly controlling the delivery of that service’s promise, runs the risk of further alienating the very partners RIT is trying to serve.
Providing Services for the Enterprise vs. Supporting Research

Service Design

From its launch, the Research Hub service has held to a set of service design decisions that make sense from an enterprise point of view. One, the service should support generalized collaboration. Two, the service should be “generally accessible ... 24/7, except for announced maintenance periods.” (From the Research Hub service level agreement, https://hub.berkeley.edu/about/ResearchHub-SLA.pdf)

These decisions have had direct ramifications:

- **Scope of responsibility, cost and ability to address needs**
  The Alfresco code base is free; its Community Edition is available for no cost. Typically, the community edition encompasses the most up-to-date enhancements and innovations. The company’s business model relies on paid licenses for its Enterprise Edition, which bring service administrators the security of greater quality assurance (due to additional engineering and formal testing) and technical support.

  Intending the Research Hub to provide storage and sharing of documents important to the daily scholarship and administrative work of anyone on campus, the service opted to go with the enterprise version. The license has cost between $25,000-$50,000 per year. Given the ensuing history with the product, featuring a number of technical concerns, this was a smart decision.

  For the campus’s Research Administration and Compliance unit, however, that wasn’t the logical decision. RAC (Neil Maxwell, Ken Geis) customized the free, community edition of Alfresco to manage the Chancellor’s correspondence [as well as communications related to research grants -- check this]. For RAC’s purposes, given the concise scope of its service and the available level of system administrator and software developer support relative to the small number of users, the cost savings outweighed the benefits of a license.

Important questions to ask, from RIT’s perspective:

- **Has RAC been more nimble in its approach to solving the needs it faced?**
  If so, is this because the scope of its service is narrow and the scale of its user base is small? Has its efforts been helped by the fact that there is no cost for the software?

- **How does the enterprise nature of a service such as Research Hub affect its responsiveness to the needs of users?** Can an enterprise service with a large, diverse user base react to changes in the campus IT ecosystem – for example, the advent of Google and Box – to change its design from “generalized collaboration” to something more specific?
Providing Services for the Enterprise vs. Supporting Research

- **Strict account management, especially de-provisioning of accounts**

Given its mandate to provide secure storage to all campus comers, Research Hub chose to tie its authentication processes to campus’s LDAP service. This design enabled the service to identify users, promoting responsibility on their part. In a period when illegal downloads posed a risk to campus systems, this was an important consideration. Even more so for a service in which users would manage and share important documents: Research Hub needed to provide a mechanism for the accountability of its users. In the current era of constant threats to Internet security, integration with campus identity management and security regimes has become vital.

At the same time, Research Hub proposed to offer a sizable amount of storage quota to everyone on campus. This feature of the service aimed to remove cost as a barrier to collaboration for those reluctant to move their practice online. It was also meant to attract campus users of any of a number of similar web based services, services which campus IT personnel were being increasingly called upon to support. The generous quota raised the concern that storage of no-longer-used materials would mount over time to an unsustainable level, as students graduated, graduate students and post-docs found jobs and positions, and faculty and staff retired or followed offers elsewhere. Tying Research Hub accounts to the CalNet directory provided a way to automatically decommission accounts when users ended their affiliation with campus. As stored content would be tied to those accounts, the service could manage the disposition of the materials.

Note that there is an access component to this issue, as well as a cost component – one that confronts services such as Google, too: if I were to leave campus, documents that I have shared in Google Drive would disappear.

This design was in contrast to that of bSpace, which handled accounts locally and had no feasible means of culling dormant accounts. The design decision meant, however, that unless a solution was found, Research Hub would only be accessible to campus members and not to their colleagues elsewhere. This would be a major impediment to many of its users. (See the next section for a discussion of how Research Hub handled this.)

Questions that arise:

- **How does campus’s responsibility for the actions and security of its members affect service design today?**

- **Should the cost of storage have such impact on the design of the service?**

- **Can campus afford to “be out in front” of issues such as managing access after accounts are closed?**
Case studies: User needs and how Research Hub has attempted to meet them

Research Hub launched in Fall 2011, shaped by the service strategies and design decisions described above. Adoption has been slow but steady, promoted through online and print announcements, campus presentations, beginner- and advanced-level training workshops and consultations, and support efforts that have doubled as outreach. Reaction to the service has generally been favorable, though many individuals and groups have tried the service and found it wanting. More saliently, however, Research Hub staff has heard – early and often – articulation of a set of needs basic to this service area. How Research Hub has addressed these needs is informative and, perhaps, instructive to the venture upon which RIT now sets itself.

• Case 1: Access to colleagues outside of UC Berkeley (i.e., without CalNet IDs)

From the very beginning, prospective Research Hub users made clear that they needed a way to collaborate with colleagues from beyond UC Berkeley. Campus’s CalNet identity management team offered to develop a solution. Research Hub partnered with the CalShare service to provide use cases, design ideas, feedback, testing and, eventually, participants for the CalNet Guest Account service. A year later, the Guest Account service was operational. With only a few clicks, campus members could sponsor colleagues for guest accounts that allowed access to Research Hub and CalShare. It was (and still is) hoped that other CalNet-based services would adopt the Guest Account service, extending the benefit of the development work widely across campus.

The CalNet Guest Account service has opened up Research Hub and CalShare to colleagues from beyond UC Berkeley. However, guest accounts have not been used as extensively as anticipated. Perhaps the service is still too rough. (Problems with guest accounts are the source of an out-sized proportion of Research Hub support tickets, including one glaring failure in which a faculty member attempting to organize a scholarly conference was unable to create guest accounts automatically from a comma-separated list of 200 invitees.) Or, perhaps, in the months it took to implement CalNet Guest Accounts, many prospective users had turned elsewhere. Users of the Google and Box services can simply share files with a colleague by entering the person’s email address. No wall restricts those services to people with campus IDs. Rather, CalNet authentication co-exists with the authentication systems provided by those services – a nice bit of usability.

(Alfresco itself is moving towards a hybrid “cloud/on premise” solution that promises to simplify access while allowing individuals and the institution to maintain some control. The price for campus access to the Alfresco cloud service has not yet been determined.)
Providing Services for the Enterprise vs. Supporting Research

Questions:

*Can campus ID management systems be made more flexible and easy to use?*

*How important is it to have institutional leverage over all users of IST services?*

- **Case 2: A place to store data securely**

In the breath that followed “We must be able to collaborate with colleagues beyond campus” – or was it the breath that preceded it? – Research Hub staff heard that campus scholars needed a secure place to store and access research data. People in fields such as psychology, public health, and social welfare voiced this requirement most loudly, but they weren’t alone. Research Hub staff responded by seeking to work with the campus’s IT policy and IT security teams to gain some sort of certification to store notice-triggering data – labeled “Protection Level 2” in the campus’s emerging minimum security standards for electronic information (MSSEI).

That process is still running its course more than two years later. Security policies have been developed; standards definitions have been worked out in conjunction with campus; the security team has presented Research Hub with a checklist of eighty or so items to address; and the paradigm has shifted to one in which IT managers and campus users share responsibility for mitigating the risks associated with breach of confidentiality.

Research Hub is finally within arm’s reach of being able to offer itself as a secure place to store research data. The burden on the user will be great: they will be required to fill out forms, monitor access status and report regularly. Not insensible requirements, but a high barrier to participation. Even after Research Hub attains that status, the decision to allow researchers to store data in Research Hub will still belong largely to the campus’s institutional review board (IRB) and the Office for the Protection of Human Subjects (OPHS).

What have researchers been doing in the meantime? Apparently, they’ve been using bSpace and CalShare, systems for which approval by the IRB dates to an earlier day. Or they’ve been encrypting the files before uploading them. Or not. Each of these solutions pre-dates the advent of Research Hub. If they were in widespread use, why would the need for a secure place to store data have been voiced so vociferously?

Like strict accountability with regard to access, secure data storage seems to be an enterprise priority that doesn’t in all ways align closely with user behavior. Moreover, fulfilling the requirements of the enterprise can take
Providing Services for the Enterprise vs. Supporting Research

months and years, encompassing technical, organizational and political elements. It is hardly the domain of nimble innovation.

Yet, it is the role of IT to define and live up to best practices, and to make those practices available to all users. The Research Hub team, with its IST partners, is proud of these efforts – and accomplishments – in the area of security, as well as in account management and broad access. We have responded to user needs and pioneered approaches for other services to follow. We have kept to high standards. Demanding that of ourselves, and delivering on it, is an essential element of our relationship with campus.

Questions:

What becomes of the users’ work in the time it takes an enterprise service to develop a satisfactory solution to its users’ problems?

At what point does the enterprise service become irrelevant?

How do we establish trust even when we don’t yet have solutions?

• Case 3: Metadata handling (custom fields per user; extracting/embedding data in files; bulk update)

Here, the issues are:

- With more resources, we ought to be able to make custom fields available. But where would the demand end, and would these custom sets clutter the interface and operation of the service for all other users? This feature is most-closely aligned with the needs of the Berkeley Art Museum/Pacific Film Archive (BAM/PFA), partners in the grant-funded pilot program from which Research Hub emerged.

- For deeper changes, we’ve elected (wisely) to defer to Alfresco. However, Alfresco has been working on, and seemingly close to delivering, additional metadata and media management features for at least a year and a half now. This, too, has been a frustration to BAM/PFA, who responded to the limited number of fields available in Research Hub by embedding their metadata within their image files. Were Alfresco to enable them to extract, edit and re-embed these property values, Research Hub might finally be seen as user-friendly in their eyes.

- As to bulk updating – making the same attribute changes to a set of files in one operation – this lack within the Alfresco platform drove the Townsend Center for the Humanities away from using Research Hub. Maybe Research Hub should take a page from the CollectionSpace web apps and provide this
Providing Services for the Enterprise vs. Supporting Research

service through a secondary interface. That would require resources, however, and incur the risks inherent in allowing non-Alfresco access to the database. Moreover, programmatic access to this shared resource could encroach on the privacy that people expect from Research Hub. As an enterprise-wide service, this option has been a non-starter.

• **Case 4: Better tools (wiki, blog, calendar, etc.)**

  - Research Hub is simply waiting for Alfresco to deliver these improvements. Not much Research Hub can do here but help people live with what’s offered. Still, that’s not a happy message to tell users, especially when it makes their jobs more difficult and time-consuming.

• **Case 5: Group management**

  Alfresco provides the means to control access to folders, files and other content in the repository by role and by individual. The process of assigning permissions to individuals can get unwieldy when more than a small number of people are involved. Research Hub site managers have called for group management capabilities, a feature that is currently available only to the system administrator. Research Hub staff, in turn, has urged Alfresco to delegate this function to site managers. This has been one of the harder lessons in the limits of repackaging third-party software.

  At the same time, campus’s CalNet team has initiated a Grouper project that could solve the problem for a host of uses and services, Research Hub included. A campus Grouper service would be a broader “win.” The project’s momentum is unclear, though, so Research Hub users wait, at the mercy of IST funding priorities.

**Conclusion**

I’ve titled this essay with a stark distinction between enterprise services and focused support to better illuminate an organizational divide that has huge consequences for central IT’s ability to partner with campus scholars. I’ve looked to the experiences of Research Hub for lessons, but examples can be found in any number of IT efforts. It is not a matter of “good guys” and “bad guys”; motivations on each side are real and honest.

Winning the trust of campus researchers will be a critical factor in the success of RIT’s new initiatives. Historically, central IT has not often been seen as a strong partner by researchers on campus. Yet these researchers still look towards IST for support. Now, as we in RIT begin to engage, design and plan with them, we need to learn to recognize this divide, so that the decisions we make and the services we provide will have an odds-on chance of bridging it.