

Workshop Notes

NMI-EDIT Advanced CAMP: Scaling Secure Collaboration

June 28-29, 2007

Portland, Oregon

<http://www.educause.edu/camp073>

NOTE: This document is intended to supplement the presentations, which will be available on the [proceedings website](#) soon.

Workshop Notes by Steve Olshansky, Internet2. Comments to <steveo AT internet2 DOT edu>

Thursdays, June 28, 2007 Collaborative Applications: Scope, Terminology, and Central Questions

• RL "Bob" Morgan

Is there is a natural tension between R&E and vendors developing applications in this space? They are trying to deliver complete packages, easily installed and managed and supported - their primary goal is collaboration *within* the organization. When you try to collaborate *between* organizations they may be looking at protocols like OpenID, unless they hear about SAML-based approaches such as Shibboleth from their customers. Thus sometimes there is a natural tension between central IT and vendors about their approach. The commercial applications tend to need to be adapted/integrated to work in our environments.

Many have observed that there are difficulties in getting application and infrastructure staff to communicate and work together. There is a trend toward disintermediating users, allowing them to break out of their silos more easily
Vendor applications - many tend to want to control identities directly rather than rely on external sources. This presents integration challenges in complex environments
Interesting applications are being built at the higher layers, which need reliable infrastructure below to support them.

Collaborative Applications, Suites, and Tools

- Duffy Gillman, Computing Manager, Principal, The University of Arizona
- Chad J. Kainz, Senior Director, Academic Technologies, University of Chicago
- John F. Walsh, Associate Vice President, Enterprise Software, Indiana University

The main goal is to facilitate sharing, remove friction, bring the various subject domains together. This is hard to do within the organization, even harder between organizations...

Reference Link: <http://rice.kuali.organization/>

The Kuali Rice effort will provide an enterprise class middleware suite of integrated products that will allow both Kuali and non-Kuali applications to be built in an agile fashion, such that developers will be able to react to end-user business requirements in an efficient and productive manner, so that they can produce high quality business applications.

ASU is hoping to release their groups work (APIs) under open-source license terms, more on this to come...

Q: Collaborations that extend our (R&E) community outward - how do we represent the institutional nature of contributions out into the world?

A: In an open application environment, there is a difference in the weight given to some voices over others - some are considered by some to be more authoritative and respected. How can we extend the academic publishing model - peer review - to this new space? Is there room for blending this somehow? How can we expose this information in a way that is easy for people to utilize, and at the same time acknowledge the value/weight of the originating voices?

Q: When applications query middleware infrastructure and present information to users in a usable fashion - do application developers have the right skills to present/expose the info to users in the best possible way? What is the right model for manipulation and presentation of data?

Collaborative Applications, Suites, and Tools

- Chad La Joie, Development Manager, Georgetown University
- James W. Laney, Software Developer, Catalyst Research & Development, Learning and Scholarly Tech, University of Washington
- Caleb Racey, Webteam ISS, Newcastle University

Q: Is data persistence in the medical space an issue?

A: Much of it is image data, which tends to not be as short-lived in terms of file formats as other data formats.

Q: How does HIPAA relate to this?

A: It varies depending upon the data in question, and whether has been deidentified, but generally there seems to be a lack of clear answers to this.

Q: Is "loose coupling" something that users "get"?

Application Integration Frameworks

- Tod Jackson, Senior Technology Architect, OpenEAI Software Foundation
- Stephen Wheat, Enterprise Architect for Administrative IT and Co-Founder OpenEAI Project, University of Illinois at Urbana-Champaign

Q: How do you handle transactions that involve multiple destinations?

A: When a module processes a request/message, and if it contains a request that it cannot fulfill, it can then republish it for consumption by other modules.

Q: How do these collaborative applications we are looking at connect with OpenEAI?

A: The core methodology of defining authoritative data sources and the consumers of appropriate data would also apply here.

+++++

Campus Infrastructure and Collaborative Application Integration

- **Steven T. Carmody, IT Architect, Brown University**
- **Michael R. Gettes, Manager, Applied Middleware, Internet2**
- **Jim DeRoest, Research Channel**

Q: How is Internet2's COManage application (under development) different from MyVOCS?

A: MyVOCS uses Sympa under the covers to manage the groups, instead of an underlying directory which is more broadly scalable and extensible, and which is what COManage uses.

In the context of delegating privilege management, using the native Signet tools, a key question would seem to be "what is the locus of the authority being expressed?"

What about de-provisioning issues? After a user deletes himself, what happens to his privileges and groups?

External data sources could be pulled in and joined to expand the attributes available to be used...

+++++

The Current Landscape Panel: Requirements, Capabilities, and Opportunities

Panel, Michael Gettes, Internet2 (Moderator)

- **Chad La Joie, Georgetown**
- **Jim DeRoest, ResearchChannel**
- **Duffy Gillman, U. Arizona**
- **Jim Laney, U. Washington**
- **Steve Wheat, UIUC**
- **Bob Jackson, U. Indiana**
- **Richard Shaw, Cal State**
- **Chad Kainz, U. Chicago**
- **Steve Carmody, Brown**
- **RL "Bob" Morgan, U. Washington**
- **Caleb Racey, Newcastle University**

Q: What other kinds of capabilities, beyond Identity and Access Management, could/should be considered for future development of the COManage application under development at Internet2 (and demo'd by its developer/integrator, Michael Gettes)?

A: Workflow is of interest to many, note that Mellon and NSF are likely going to jointly sponsor a workshop on this topic in October 2007

Observations:

- A portal or some sort of container to house/manage the assembled applications would seem to be a likely direction
- OKI Core service APIs would seem to be useful here
- OSIDs, repository for shared objects

Q: Where should the integration be happening?

A: Not in the applications themselves, but rather between them. What advice/direction can be given to app developers to more easily tie into this infrastructure?

Q: How scalable are these applications, really? When these applications are integrated, are there collective effects of scale that come into play?

A: Even well-funded science projects may not have the resources to fully develop these applications. The deploying organizations have a role here, as they integrate these applications into their infrastructure.

Why isn't there something like a simple hosted collaborative space, like an "eScience SourceForge"? Is there a business case for it? A number of scientists have been talking about this, but their funding usually does not cover this sort of work (i.e. collaboration tools and infrastructure) and thus they are looking for other ways to support and develop this kind of work.

Is a classic business case even relevant here? If the scientists *really* wanted to do this wouldn't they find a way to make it happen - especially in light of all of the other collaborative applications springing up with no viable business model? The Grids community has a great deal of funding and thus presumably they could do something like this if they found it important, whereas the open source community in general is not so well funded.

Do "problems" fall into the category of dissemination, or collaboration? The core publishing model drives some of the thinking around this.

Q: What about leveraging some user databases out in the wild, such as in the medical space? Or are there corollaries to AAMC (Association of American Medical Colleges - aamc.org) in e.g. the Physics space?

Q: How do we keep the *entire* higher ed community up to date in terms of cyberinfrastructure, given continual funding challenges? How can we make the middleware infrastructure more accessible to smaller institutions, with their attendant challenges in staff and funding?

Is there a role here for middleware service providers? Probably yes, but they need the right kind of middleware to provide - "lean," easy to integrate, etc. What role could consortia play here? Lead campuses in a state system? Consulting firms?

Open source software is becoming more accepted, there is more support available (from the user/developer communities and from consulting and integration firms) and the overall quality is getting to the point where it is often very competitive with commercial applications.

Smaller schools have tended to gravitate to Microsoft solutions because they are so staff/\$ constrained that they really need to go with what is easiest to use, deploy, and support.

How can we as a community work with the software vendors to help them understand that standards compliance is important to us as a significant market, and thus to them. In turn this makes their products better and easier to use. Many of the larger vendors tend to take the easy way out based on what they perceive to be the needs of the smaller schools, rather than taking a more expansive view of the opportunities available to them.

=====

Friday, June 29, 2007

Campus Infrastructure and Developers Perspective

- **Wilson D'Souza, Director - Infrastructure SW & Architecture, MIT**

I am not interested in reinventing alternatives to good options for various collaboration tools (e.g. wikis, blogs, Subversion, JIRA, etc.) already available, but rather in aggregating them into a coherent interface that takes advantage of their respective strengths and presents a usable suite to users. In a sense this can be viewed as a large mashup.

MIT has committed serious resources to moving this along, including offering the assistance of an architect, as well as integration help, at no cost to various departmental projects if they use our supported tools. If they choose other options they are on their own. We don't force them to use our supported tools if they want to use alternatives, but it is discouraged. Most appreciate the value in using our supported tools and the support we are willing to provide them.

What is the security architecture for all of this? It varies depending on the specific component, but there is an overall security framework/strategy in place that tries to be flexible.

There are some departments with special needs which want to run some of their own instances of these services (e.g. because of DoD work), and central IT will support them with the understanding that they will utilize the same suite of components that have been integrated elsewhere on campus, to the extent practical for their needs.

=====

A Focus on Policy and Process Issues

- **Margaret Harrington, Director Organization Improvement Services, University of Southern California**
- **George Mathew, Chief Technology Architect, Fox Chase Cancer Center**
- **David H. Walker, Director of Advanced Technology, University of California Office of the President**

Common elements of successful collaboration:

- Common vocabulary
- Governance
- Credentialing
- Auditing
- Consensus/buy-in
- Convenience v. risk
- Many projects cross organization boundaries, and some cross institutional boundaries
- Synchronous and asynchronous collaboration
- Complexity is a fact of life, and must be embraced for successful collaboration.

"Academic cyberinfrastructure" - set of services layered on top of federated IdM infrastructure

=====

Conclusions and Possible Directions

- **Thomas J. Barton, Senior Director for Integration, University of Chicago**
- **R.L. "Bob" Morgan, Senior Technology Architect, University of Washington**

In the real world, users generally have multiple affiliations, wearing different hats in different contexts. This unavoidable complexity must be addressed in any successful collaboration system. How can these various relationships be accommodated, reducing the friction as much as practical?

Q: Registration model? Invitation model?

A: This tends to fall apart for users in a federated model, since users doing the inviting are doing so with a specific affiliation. Does this then imply a role for an outside Identity Provider (IdP) to serve as a separate layer to facilitate this?

The model only works for systems in which your identity proofing is based on e-mail - i.e. who receives the invitation e-mail?

Q: How does a large scale federated people-picker work in this context?

Q: Are human-understandable identifiers a viable option? Seemingly not so far.

Complexity is a fact of life, but whenever possible when designing applications it makes a lot of sense to try to address the simpler (lighter weight) use cases first, since they may end up being sufficient for most of the real world use cases. Then if needed, additional complexity may be added down the road...

Q: How can general-purpose tools be modified so as to be more usable for specific needs? Reinventing the wheel to recreate tools is not a practical approach.

Management, policies, and procedures are ultimately more important and impactful than the particular technology(ies) utilized. Some folks may tend to get too focused on a particular set of tools, it is helpful to try to raise the conversation up a level to get to the real needs to be addressed and goals to be achieved.

More loosely coupled, less integrated approaches are often very useful incremental approaches, and have the added benefit of allowing modification midstream as experience is gained and feedback received.

When users collaborate in the real world, loss of "control" often goes with the territory. What does this mean for domain-specific policies/procedures /compliance?