

Conference Call Minutes, 2008-02-07

ITANA Conference Call Minutes

February 7, 2008

****Attendees****

Jim Phelps, University of Wisconsin - Madison
Piet Niederhausen, Georgetown University
Tom Barton, University of Chicago
Ron Thielen, University of Chicago
Klara Jelinkova, Duke University
Chris Phillips, University of Maryland
Jon Giltner, University of Colorado
David Walker, University of California Office of the President
Steve Mullins, University of Alaska
Steve Kellogg, Penn State University
George Brett, Internet2
Steve Olshansky, Internet2
Dean Woodbeck, Internet2 (scribe)

****Agenda****

(0) Roll Call. Agenda Bash.

1. Accept minutes of last call
2. Data Classification and Data Retention Policies - Klara J
3. The Open Group's EA Practitioners Summit Report Out - Jim P

****EA Practitioners Summit****

Jim Phelps discussed the Open Group's EA Practitioners Summit he attended last week. The discussions were at a high level and didn't bring things down to the detail level very well. Most of the conversations were similar to those on this phone call - trying to get alignment, getting groups to think at a higher level, and a lot of talk about defining Enterprise Architecture. It reinforced his thoughts about EA - that is, the enterprise architect's role is amorphous and has to fit the institution's culture and governance. He would not recommend this meeting unless it were held very close to your campus.

****Data Management****

Klara sent several example use cases to the email list in preparation for today's meeting. The question at hand is what to do about managing enterprise data, in light of such areas as retention, legal requirements, end user needs, and the university's policies for online archives.

The example use cases include:

1. Online archives of university functions - Duke Chapel, recordings of events.
2. Multimedia archives of class support materials (not just lectures but also study materials, etc)
3. Images as class materials but also products of artist's work (art history as well visual studies)
4. Research data - as it pertains to new NIH rules
5. Research data as institutional asset housed at the institution
6. Research data of Duke Pls at other institutions or on Google docs or in the "computing cloud"
7. E-mails as electronic archives under e-discovery rules
8. Logs as electronic archives of access to systems and actions taken.

There are many considerations as we start to tackle such projects. In some cases, the data may be in its native application and would need to be translated into XML or some other enterprise data format to be usable by other systems. There is also the difference between in-place retention vs. reflecting the data in a data warehouse, for example.

The working group discussed a number of issues related to data retention and data retention policies.

The University of Colorado has policies around information assets within departments. Assets can be classified as private, semi-private or public; or, along a different dimension, whether it is critical or not critical, in terms of business function. Another layer involves the archival policies for different types of data, such as student academic information or email, for example.

There was a general discussion about the role of the library and the relationship between IT, the library and data retention. It was generally agreed that libraries take a broader view of data. Their framework is more about collections and managing a collection over a period of time. They have a long-term point of view. They also have the skills and resources to do cataloging and make the data searchable.

When data is unstructured, such as raw data that doesn't seem to fit with a larger collection, libraries, in general, don't have the interest or motivation to do the management and preservation. It may be important to archive some institutional data, just for the sake of archiving it, but who has responsibility for that? And can the library and IT work to help educate end users on what is worth preserving?

Decentralization of data control raises another dimension to this discussion. Some faculty, for example, are using such external services as Flickr and Google Docs for storing course content. This raises the question of whether the data is owned by faculty or the institution. If it is the institution, what are the data retention issues to using third-party providers for storage.

The conversation then came full circle asking how IT can assist the institution in having a conversation about institutional data management.

The University of Chicago has created a program of data stewards and data managers. The stewards have policy level ownership of data and the managers are the people who understand the data at the field level. We are going to try to have these conversations with these people. The University of Maryland has a similar program

This may work with institutional data, but still leaves open the question about research data - dealing with individual researchers and data that may or may not be maintained at the institution.

One strategy to move this conversation might be to develop strategies in smaller situations, like individual use cases, rather than trying to tackle all of these issues institution-wide.

As a next step, Klara will enter her use cases in a spreadsheet, listing the stakeholders, data stewards and data managers. This will be placed on the wiki, inviting ITANA members to enter their own data and, hopefully, recognize a use case that they have explored. Using specific areas, perhaps we can gather feedback on use cases and experiences.

In addition, Jim asked that anyone with policies in any of these areas to place them on the wiki, or place a link on the wiki back to the policies.

Jim will also send a reminder to the ITANA list reminding people how to gain edit access to the wiki.

With all of this discussion and gathering of information, we should be looking down the road to a publication or presentations based on what we learn. That would help to get the information out to larger audiences.

It was reported that the last CSG (Common Solutions Group) meeting included a round-up of future topics and data retention was one.

Klara will be doing a workshop at the CSG cyberinfrastructure workshop, May 14-16. She is wondering if there might be enough information gathered to include in her presentation.

****Future Agenda****

Jim reported on a document that was released having to do with Mellon enterprise service bus work and will invite someone to speak on this during a future call.

Herbert's larger enterprise information management survey will be a topic on a future call.

If others have items for future agendas, please contact Jim.

****Upcoming Meetings****

- ITANA members reported they will be attending these upcoming events:
- Internet2 CAMP: Bridging Security and Identity Management in Tempe, Feb. 13-15
- EDUCAUSE Midwest regional in Chicago, March 17-19
- EDUCAUSE Enterprise Information and Technology Conference 2008, Chicago, May 28-29
- Department of Defense Enterprise Architecture conference, Orlando, April 14-18.

An ITANA face-to-face will take place in June, in conjunction with an advanced CAMP. The location is to be determined.

Those attending conferences are asked to provide a brief report back to ITANA on your experience.

****Next Call - Thursday, February 21, 2008 - 2:00 p.m. EST****