Fostering LIMS Development Through Open Standards

Part II – Ontologies and Business Process

Avrum Goodblatt

PathBioResource, U. PENN School of Medicine

goodblat@mail.med.upenn.edu

This is the second in the series of short articles about Laboratory Information Management Systems (LIMS). This installment will focus on two issues – looking at ontologies and at business process design. The goal is to demonstrate how LIMS is really a combination of several capabilities, and that although each capability should be looked at separately, they ultimately must all work together as seamlessly as possible.

Before diving in, I would like to thank Dr. Q. C. Yu, the director of our Biomedical Imaging Facility here in Pathology and Laboratory Medicine at Penn http://www.med.upenn.edu/bmcrc/morph/?morph for his assistance, encouragement, and sponsorship of my participation in the Honolulu conference.

I would also like to thank the people who have written in and described the work they are doing. I neglected to ask permission to repeat the letters so I have not quoted anyone here from these sources. In the future, please assume that anything sent to me might appear in these articles. The mailing list is still not up, mostly due to my busy schedule, but look for it in January.

As I described previously, there are four critical goals:

1. Capture billing, user, and work description information (using a shared ontology).
2. Integrate scheduling and logon to equipment using institution-level authentication.
3. Secure file server with access tied into the service request.
4. Generate logs and invoices automatically and if possible feed into the institution billing system.

An open source intiative that recently came to my attention is the Open Microscopy Environment. "http://www.openmicroscopy.org/" The OME is described on their website at: "http://www.openmicroscopy.org/concepts/index.html." Its major purpose is to maintain a database of images with annotations, thus this OME is concerned extensively with the ontology. It is not a complete LIMS, according to the list above, but is focusing on the central issue of how to store, annotate, and retrieve images regardless of their derivation. Since it is open source, it is likely that it could be relatively easily linked into the other modules in our list. The following is a quote from Jason Swedlow at OME:

"Biological microscopy has always required an "imaging" capability: traditionally, an image of a sample was drawn on paper, or with the advent of light-sensitive film, recorded on media that conveniently allowed reproduction. However, the application of digital detectors to microscopy has converted the biological microscope into an assay device that linearly measures the flux of light at different points in a cell or tissue. The transition of a microscope's output from "image" to "data" has created new demands for storage, analysis and visualization that are not adequately met in any available software package. This absence of suitable software for image management currently hinders many projects from exploiting the full potential of digital microscopy to solve biological problems. For example, it is a major stumbling block

for cell biologists wishing to use the powerful techniques now available for performing live cell dynamics, photobleaching and FRET studies on cells expressing fluorescent protein fusions. In addition, cell-based high content assays are under development in many academic and commercial labs, but there are no tools available for managing this type of data and integrating all experimental information and data analysis. Overcoming these difficulties will therefore have an immediate and valuable impact on many areas of cell biology.

To solve this problem, we are building an open source image data, metadata, and analysis management system known as the Open Microscopy Environment (OME; http://openmicroscopy.org). OME is specifically designed for the storage and analysis of digital microscope image data and metadata. The major focus of OME is not on creating novel analysis algorithms, but instead on development of a structure that ultimately allows any application to read and use any data associated with or generated from digital imaging microscopes."

The OME ontology, which is currently geared toward cellular dynamics and cellular localization or phenotypes, could be expanded to other imaging areas. I will report more at length on this in future articles.

One of the benefits to open source is that it usually attempts to adhere to open standards (using open standards however does not require using open source). Open standards are useful in that they facilitate software interoperability. For example, in today's microscopy world, different vendors use different formats for their images, and it is not always possible for one set of software to read another vendor's images. Open standards, however, apply to more than images. They can specify how a set of software talks to a microscope or camera, or how the acquisition software talks to the scheduling utility. As more open standards are developed and adhered to, the more likely it is that a complete LIMS can be built.

The scheduling and tracking features of a LIMS system especially lend themselves to open standards. We are attempting to offer our researchers and facilities staff — and the relevant PI's and Business Administrators — the ability to see at a glance various stages of their work. We have defined the following stages:

- Request
- Approval
- Receipt
- Assignment
- Quality Control
- Execution
- Delivery of results (upload)

Some stages are of more importance to the staff than to the investigators, and there are certainly more stages that could be enumerated for the facility staff. The point is that this sort of work tracking is handled well by workflow management software. One group working in this area is the WfMC (workflow management coalition) http://www.wfmc.org/index.html

While the WfMC works with all sorts of business process foci, much of which would appear to be far removed from a microscopy facility, the WfMC efforts to create open standards have encouraged the development of open source standards-based workflow software, such as that found at http://www.enhydra.org . These products can then be used in conjunction with our LIMS.

I have not yet used these tools myself, but it appears one can design a workflow using visual tools, and then have it translated into a program to manage that workflow. Then, all the different LIMS users at your institution would have to do would be to tap into this

workflow server to find out what they are supposed to be doing next, or report on work they have done. The same server could then report all of this to the relevant people as mentioned above.

It would be nice to be able to just get a system, which had all these components put together. That will never be possible — since one of the components is billing, and that needs to be customized to the billing system already in place in the institution where the facility is housed. We are now working on the piece to hook our invoice generation system into our universities journaling operation (for internal clients). Others may not have such access to their institutional system, but still I would prefer that the folks who make billing systems not try to build LIMS. What LIMS vendors need to keep in mind only is that they should provide open standard entries to allow their LIMS to generate what is needed for the billing and journaling.
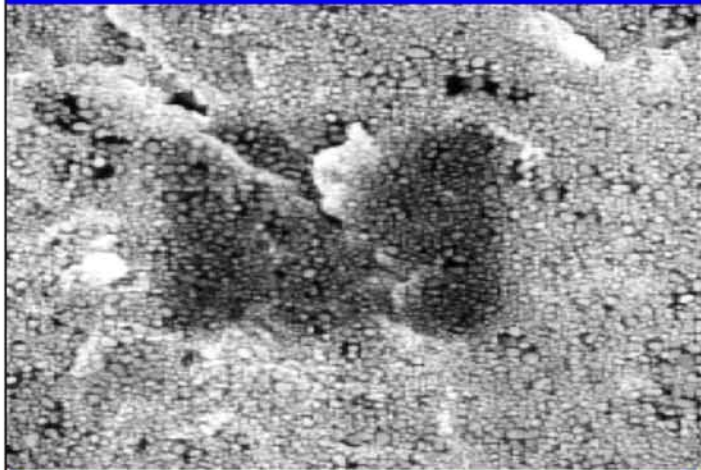
Another issue that has come up is how to limit access. We have come up with three models — a minimalist, a maximalist, and a compromise.

- Minimalist allows anyone access to anything. Some smaller operations have this de facto.
- Maximalist says that only the researcher and the PI paying for the research has access.
- The Compromise is what we hope to provide in the near future – the researcher and anyone in the researcher's lab has access. We found that most labs find this sufficient, but that is probably not going to be true in a corporate environment.

It is important when choosing a LIMS to find out if its authentication and access capability can hook into your existing model. For example, we are trying out a LIMS for another facility here at Penn that does not have that ability at a sufficient level. It can hook into our user model, but not into our lab and project listings. There is no easy way to get the billing information into the experiment's metadata. This is also true of several of the microscopy packages I examined at the M&M-2005 meeting. We are advocating a university-wide model for tracking project authorization. If anyone is at a university where something like that is available, please let us know. In the meantime, we have taken the approach of sharing our authorization system with other departments, with some success. Perhaps one day we can even have the various microscopy facilities share this sort of back-end data.

As I mentioned at the top, I will get the mailing list operational. One of the first activities could be to have reviews of LIMS, as well as links to descriptions (and downloads?) of existing microscopy LIMS. If any one wants to send me links I will definitely put up a page with the proper attribution. Vendor email is welcome too. I am especially interested in hearing from those who are familiar with the challenges of connecting vendor camera/microscope software to existing authentication systems and databases. Please also let me know which of these issues is of most interest and I will devote future articles to those matters. Suggestions for LIMS operations to review would also be welcome, as would be suggestions for standards to submit to the MSA for review and (we hope) endorsement. Finally, if you are getting questions from your administration about LIMS and would like help getting them answers, please email the questions to me — I will share our experience here at Penn and post them on the list. After all, this stuff is not free (if not the software, then the time must be allocated), and we all could use help in preparing the justifications and budgetary requirements and time lines involved. ∎