



# *Institute for Laboratory Automation*

*Realizing the Potential of Automation & Information Technologies in Science*

## **Integration in Laboratory Automation & Informatics**

### *How do we get there?*

Joseph G. Liscouski

Institute for Laboratory Automation, (a non-profit organization)

Groton, MA

Web site: <http://www.InstituteLabAuto.org>

Email: [j.liscouski@InstituteLabAuto.org](mailto:j.liscouski@InstituteLabAuto.org)

This document was developed with the intent of presenting ideas and eliciting comments. Based on the comments received, we'll determine our next course of action.

In John Trigg's weblog "The Integrated Lab"<sup>1</sup> there have been a number of discussions about the need for integration in laboratory systems as well as questions about what it will take to get us to the point where integration is possible. The point of this article is to provide one answer to those questions.

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

[www.InstituteLabAuto.org](http://www.InstituteLabAuto.org)  
[info@InstituteLabAuto.org](mailto:info@InstituteLabAuto.org)  
a non-profit organization

© 2009 Institute for Laboratory Automation, Groton, MA, All Rights Reserved

The purpose of *integration* in the lab is to make it easier to connect systems; for example, pass results from a chromatography data system to a Laboratory Information Management System (LIMS) or Electronic Laboratory Notebook (ELN) and then on to other groups. The results of this ability to integrate systems include:

The purpose of *integration* in the lab is to make it easier to connect systems.

- Smoother workflow – less manual effort, avoiding duplication of data entry, this is something that people are striving for and accomplishing in production environments including manufacturing, video production and graphics design,
- Easier path for meeting regulatory requirements – integrated systems, with integration built in by vendors, results in systems that are easier to validate and maintain,
- Reduced cost of development and support,
- Reduction in duplication of records, better data management,
- More flexibility – as we’ll discuss below, integrated systems built on modular components will make it easier to upgrade/update systems, and meet changing requirements.

The inability to integrate systems and components through vendor provided mechanisms results in higher development and support costs, increased regulatory burden, and reduced likelihood that projects will be successful.

#### *What is an Integrated System in the laboratory?*

Phrases like “integrated system” are used so commonly that it seems as though there should be and instant recognition of what they are. The words may bring a concept to mind, but do we have the same concept in mind? For the sake of this discussion that concept has the following characteristics:

- A given piece of information is entered once, and then is available throughout the system, restricted only by access privileges. The word “system” is the summation of all the information handling equipment in the lab. It may extend beyond the lab if process connections to other departments are needed.
- The movement of materials (sample prep for example) and data / information is continuous from the start of a process through to the end of that process without the need for human effort. The sequence doesn’t have to wait for someone to do a manual portion of the process in order for it to continue, aside from policy conditions that require checks, reviews, and approvals before subsequent steps are taken.

An integrated system should result in a better place for people to work. People wouldn’t be used as a means of achieving repetitive work. Doing so results in two problems: people get bored and make mistakes (some minor, some not, both of which contribute to variability in results), and, the progress of work (productivity) is dependant on human effort which may limit the number of hours that a process can operate. It is also a bad way of using intelligent, trained personnel.

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

www.InstituteLabAuto.org  
info@InstituteLabAuto.org  
a non-profit organization

## *A Brief Historical Note*

For those who are new to the field, we've been at this for a long time, with not as much to show for it as we'd expect, particularly when compared to other fields that have seen an infusion of computer technologies. During the 1980's the pages of *Analytical Chemistry*<sup>2</sup> saw the initial ideas that would shape the development of automation in chemistry. Ray Dessy's article on LIMS, robotics, networking, and instrument data systems, laid out the promise and expectation for electronic systems used to acquire and manage the flow of data and information throughout the lab.

That concept - the computer integrated lab - was the basis of work by instrument and computer vendors, resulting in proof-of-concept displays and exhibits at PITTCON<sup>3</sup> and other trade shows. After 30+ years we are still waiting for that potential to be realized, and we may not be much closer today than we were then.

What we have seen is an increase in the sophistication of the tools available for lab work, client-server chromatography systems, electronic lab notebooks – in their varied forms, are just two. In each case we keep running into the same problem, the ability to connect things into working systems. The result is the use of product specific code, work-arounds to moving and parsing data streams. These are fixes, not solutions. Solutions require careful design not just for the short-term-what-do-we-need-today but long term robust designs that permit graceful upgrades and improvements without the need to start over from scratch.

## *The Cost of the Lack of Progress*

Every day the scientists and technicians in your labs are working to produce the knowledge, information, and data [ K / I / D ] your company depends upon to meet its goals.

*Millions of dollars have been spent on failed systems.*

That K / I / D is recorded in notebooks and electronic systems. How well are those systems going to support your need for access today, tomorrow, or over the next 20+ years? This is the minimum most companies require for guaranteed access to data.

The systems being put in place to manage laboratory K / I / D are complex. Most lab data management systems (Laboratory Information Management Systems -LIMS, Electronic Lab Notebooks - ELNs, and some instrument data systems) are a combination of four separate products: hardware, operating system, database management system, and the application you and your staff uses. Each from a different company, each with its own product life cycle. Which means that changes can occur at any of those levels, asynchronously, without consideration for the impact they have your ability to work.

Lab managers are usually trained in the sciences and personnel aspects of laboratory management. They are rarely trained in technology management and planning for laboratory robotics and informatics – the tools used today to get laboratory work done and manage the results. The consequences of inadequate planning can be significant:

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

“In January 2006, the FBI ended the LIMS project, and in March 2006 the FBI and [the vendor] agreed to terminate the contract for the convenience of the government. The FBI agreed to pay a settlement of \$523,932 to the company in addition to the money already spent on developing the system

www.InstituteLabAuto.org  
info@InstituteLabAuto.org  
a non-profit organization

and obtaining hardware. Therefore, the FBI spent a total of \$1,380,151 on the project. With only the hardware usable, the FBI lost \$1,175,015 on the unsuccessful LIMS project.” OIG Audit Report 06-33

Other examples of problems in projects:

- A “2006 ALA Survey on Industrial Laboratory Automation” published in the August 2007 edition of the Journal of the Association for Laboratory Automation posed the following question: My Company / Organization’s Senior Management Feels its Investment in Laboratory Automation Has: Succeeded in delivering the expected benefits (56%), produced mixed results (43%), has not delivered the expected benefits (1%). *44% failed to fully realize expectations.*
- "As the statistics show that 60% of all LIMS projects have failed to have a 100% successful go-live.." and this is reported in: [http://www.scientific-computing.com/features/feature.php?feature\\_id=132](http://www.scientific-computing.com/features/feature.php?feature_id=132).
- The Standish Report on project failures (looking at Enterprise Resource Planning [ERP] implementations – similar in scope to LIMS & Electronic Lab Notebooks) shows that over half will fail, 31.1% of projects will be canceled before they ever get completed. Further results indicate 52.7% of projects will cost over 189% of their original estimates.

We’ve received a number of emails discussing the results of improperly managing projects, all have requested that any identification be kept confidential. Among them are:

- A LIMS customer was given a precise fixed-price quote somewhere around \$90,000 and then got hit with several \$100,000 in extras after the contract was signed – contributed anonymously.
- A major pharma company some years back that implemented a LIMS with a lot of customization that was generally considered to be successful, until it came time to upgrade. They couldn’t do it, and went back to square one and purchased another system -- contributed anonymously.
- Reports of robotics system failures totaling over \$700,000.
- Examples where vendors are using customer sites as test-beds for software development.
- A set of 3 labs that were trying to use the same system [to reduce cost] for different types of labs with different requirements - \$500,000 spent before the project was cancelled.

In addition to those costs, there are the costs of missed opportunities, project delays, departmental & employee frustration, and the fact that the problems you wanted to solve are still sitting there.

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

[www.InstituteLabAuto.org](http://www.InstituteLabAuto.org)  
[info@InstituteLabAuto.org](mailto:info@InstituteLabAuto.org)  
a non-profit organization

© 2009 Institute for Laboratory Automation, Groton, MA, All Rights Reserved

Page 4

The causes for failures are varied, but most include factors that could have been avoided by making sure those involved were properly trained:

*The causes for failures are varied...*

- poor planning, unrealistic goals (in part because the features needed to make systems work together aren't there)
- inadequate specifications, including regulatory compliance requirements
- project management problems
- scope creep
- lack of experienced resources

The lack of features that permit the easy development of integrated systems can also be added to that list. That missing element causes projects to balloon in scope, requiring people to take on work that they may not be properly prepared for, or projects that are not technically feasible, something developers don't realize until they are deeply involved in the work.

The method people use to achieve integration today results in cost overruns, project failures and systems that can't be upgraded or modified without significant risk of damaging the integrity of the existing system. One individual reported that his company's survey of customers found that systems were integrated in ways that prevented upgrades or updates; the coding was specific to a particular version of software, and any changes could result in scraping the current system and starting over.

One way of achieving "integration" is similar to integrating household wiring by hard-wiring all the lamps, appliances, TV's, etc. to the electrical cables. Everything is integrated, but change isn't possible without shutting off the power to everything, going to the wall and making the wiring changes, and then repairing the walls and turning things back on. When we think of integrating systems, that's not the model we're considering – but from the comments we've received, it is the way people are implementing software. We're looking for the ability to connect things in ways that permit change, like the wiring in most households – plug and un-plug.

That level of compatibility and integration results from the development of standards for power distribution and for connections: the design of plugs and sockets for specific voltages, phasing, and polarity so that the right type of power is supplied to the right devices. There are other examples of the ability to connect systems:

- Universal Serial Bus (USB) – the same connector can be used to connect a computer to storage, a camera, scanner, printer, communications, etc.
- Modular telephone jacks and tone dialing allow for the telephone systems we have today – we couldn't have the level of sophistication we have now if we relied on rotary dials and hardwired phones.

These are just a few examples of component connections that can lead to systems integration. When we consider integrating systems in the lab, we need to look at

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

www.InstituteLabAuto.org  
info@InstituteLabAuto.org  
a non-profit organization

connectivity and modularity (allowing us to make changes without tearing the entire system apart) as goals.

### *What do we need to build integrated systems?*

The lab systems we have today are not built for integration system wide. They are built by vendors and developers to accomplish a set of tasks, connections to other systems is either not considered, or avoided for competitive reasons. If we want to consider the possibility of building integrated systems there are at least five elements that are needed:

1. Education
2. User Community Commitment
3. Standards – file format and messaging / interconnect
4. Modular Components
5. Stable Operating System Environment

## **1. Education**

Facilities with integrated systems are built by people trained to do it. This has been discussed within the concepts of *Laboratory Automation Engineering* published in the Journal of the Association for Laboratory Automation<sup>4</sup>.

The educational issues don't stop there. Laboratory management needs to understand their role in technology management. It isn't enough to understand the science and how to manage people as was the case thirty or forty years ago. Managers have to understand how the work gets done and the technology used to do it. The effective use/miss-use of technologies can have as big an impact on productivity as anything else. The science also has to be adjusted for advanced lab technologies. Method development should be done with an eye toward method execution – can this technique be automated?

## **2. User Community Commitment**

Vendors and developers aren't going to provide the facilities needed for integration unless the user community demands them. Suppliers are going to have to spend resources in order to meet the demands for integration and they aren't going to do this unless there is a clear market need and users force them to meet that need. If we continue with “business as usual” practices of force fitting things together and not being satisfied with the result, where is the incentive for vendors to spend development money? The choices come down to these: you only purchase products that meet your needs for integration, you spend resources trying to integrate systems that aren't designed for it, or, your labs continue to operate as they have for the last 30 years – with incremental improvements.

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

www.InstituteLabAuto.org  
info@InstituteLabAuto.org  
a non-profit organization

### 3. Standards

Building systems that can be integrated depend upon two elements in particular: standardized file formats and messaging/interconnect systems that permit one vendor's software package to communicate with another's.

#### File Format Standards

The output of an instrument should be packaged in an industry standardized file format that allows it to be used with any appropriate application. The structure of that file format should be published and include the instrument output plus other relevant information such as date, time, instrument ID, sample ID read via barcode or other mechanism, instrument parameters, etc. Digital cameras have a similar set up for their raw data files: the pixel data and the camera metadata that tells you everything about the camera used to take the shot.

In the 1990's the Analytical Instrument Association (now the Analytical and Life Science Systems Association) had a program underway to develop a set of standards for chromatography and mass spectrometry. The program made progress and was turned over to the ASTM where it is currently stalled. It was a good first attempt. There were several problems with it that bear noting. The *first* point is found in the name of the standard – *Analytical Data Interchange Standard*. It was viewed as a means of transferring data between instrument systems, and served as a secondary file format, with the instrument vendors being the primary format. This has regulatory implications since the FDA requires storage of the primary data and that the primary data is used to support submissions. It also means that files have to be converted between formats as it moves between systems.

Ideally, the standard format would be THE format for an instrumental technique. Data collected from an instrument would be in that format and be implemented and used by each vendor. In fact, it would be feasible to have a circuit board in an instrument that would function as a network node. It would collect and store instrument data and forward it to another computer for long-term storage, analysis and reporting, thus separating data collection and use. A similar situation currently exists with instrument vendors that use networked data collection modules. The issue is further complicated by the nature of analytical work. A data file is meaningless without its associated reference material: standards, calibration files, etc., that are used to develop calibration curves and evaluate qualitative and quantitative results. While file format standards are essential, so is a second-order description: sample set descriptor's that provides a context for each sample's data file – a sample set might be a sample tray in an autosampler, the descriptor would be a list of the tray's contents (standards, sample ID, etc.).

The *second* issue with the AIA's program was that it was vendor driven with little user participation. The transfer to the ASTM should have resolved this but by that point user interest waned. People had to buy systems and they couldn't wait for standards to be developed and implemented. The transition from proprietary file formats to standardized formats has to be addressed in any standards program.

The *third* issue is standards testing. Before you ask a customer to commit their work to a vendor's implementation of a standard, they should have the assurance, through an independent third-party, that things work as expected.

## Messaging / Interconnect Standards

Developers and vendors design programs to be self-standing: the software works as though nothing else existed and it is self-sufficient for all critical tasks. That is a reasonable viewpoint since it may in fact be true; there isn't any standard suite of lab software. It is also true that software exists and functions in concert with other programs and they may have the need to exchange data elements. We need a standard for inter-task communications. The advent of ELNs only raises the level of complexity. Files can be imported / exported, but if we want integration we need communications between elements. That includes the modules used in sample preparation as well as large instrument data systems, LIMS, and ELNs.

## 4. Modular Systems

The paragraph above notes that vendors have to assume that their software may be running in a stand-alone environment in order to ensure that all of the needed facilities are available to meet the users needs. This can lead to duplication of functions. A multi-user instrument data system and a LIMS both have a need for sample login. If both systems exist in the lab, you'll have two sample login systems. The issue can be compounded beyond that with the addition of more multi-instrument packages.

Why not breakdown the functionality in a lab and use one sample login module? It is simply a multi-user database system. If we were to do a functional analysis of the elements needed in a lab with an eye toward eliminating redundancy and duplication, designing components as modules, integration would be a simpler issue.

A modular approach – login module, lab management module, modules for data acquisition, chromatographic analysis, spectra analysis, etc. – would provide a more streamlined design, with the ability to upgrade functionality as needed. For example, a new approach to chromatographic peak detection, peak deconvolution, could be integrated into an analysis method without having to reconstruct the entire data system.

When people talk about modular applications, the phrase "LEGO® like implementation" comes up. It is a good illustration of what we'd like to accomplish. The easily connectable blocks and components can be structured in a wide variety of items. All based on a simple *standardized* connection concept. There are two differences that we need to understand. In LEGOs almost everything connects.. in the lab connections need to make sense. Secondly LEGOs is a single-vendor solution; unless you're THE vendor that isn't a good model. A LEGOs-like multi-source model (including open source) of well-structured, well-designed and supported modules that could be connected / configured by the user would be an interesting approach to the development of integratable systems.

Modularity would also be of benefit when upgrade / updating systems. With more functions distributed over several modules, the amount of testing & validation needed would be reduced. It should also be easier to add functionality. This isn't fantasy, this is what systems engineering – *Laboratory Automation Engineering* – is when you look at the entire lab environment rather than implementing products task-by-task in isolation.

Institute for  
Laboratory  
Automation  
Groton, MA  
978.449.7139

www.InstituteLabAuto.org  
info@InstituteLabAuto.org  
a non-profit organization

## 5. Stable Operating System Environment

The foundation of an integrated system must be a stable operating environment. Operating system upgrades that require changes in applications coding are disruptive and lead to a loss of performance and integrity. It may be necessary to forgo the bells-and-whistles of some commercial operating systems in favor of open source software that provides required stability. Upgrades should be improvements in quality and functionality where that change in functionality has a clear benefit to the user.

### *Where do we go from here?*

As noted in the initial paragraph the point of this note is to move the discussion from wanting integration to what it will take to get there. The items noted above are just introductory comments, each could be a healthy document by itself. At some point these steps are going to have to be taken. Until they are, and they result in tools you can use, labs – your labs – are going to be committing the results of your work into products and formats you have little control over. That should not be an acceptable situation; the use of proprietary file formats that limit your ability to work with your company’s data should end and be replaced with industry standard formats that give you the flexibility to work as you choose, with whatever products you need.

We need to be very deliberate in how we approach this problem. In the file format standards discussion for example, it was noted that the data file for a single sample is useless by itself. If you had the file for a chromatogram for instance, you could display it, look at the conditions used to collect it, but interpretation requires data from other files, so standards for file sets have to be developed. That wasn’t a consideration in the original AIA work on chromatography and mass spec (though it was in work done on ICP standards for the Army Corp of Engineers).

The first step in this process is for lab managers and IT professionals to become educated in laboratory automation and what it takes to get the job done. The role of management can’t be understated, they have to sign off on the direction work takes and support it for the long haul. The education needs to focus on the management and implementation of automation technologies, not just the underlying science; it is the exclusive focus on the science that leads to the smokestack / silo implementations we have today. The user communities active participation in the process is central to success, and unless that group is educated in the work, the effect of that participation will be limited.

Secondly, we need to renew the development of industry standard file formats, not just from the standpoint of encapsulating data files, but formats that ensure that the data is usable. The initial focus for each technique needs to be a review of how laboratory data is used, particularly with the advent of hyphenated techniques, and use that review as a basis for defining the layers of standards needed to develop a useable product. This is a complex undertaking, but worth the effort. If you’re not sure, consider how much your lab’s data is worth and the impact of its loss.

In the short term we need to start pushing vendors – you have the buying power – to

develop products with the characteristics needed to allow you to work with and control the results of your lab's work. Products need to be developed to meet your needs, not the vendors/developers. Product criteria needs to be set with the points above in mind, not on a company-by-company basis but as a community; you're more likely to get results with a community effort.

Overcoming the barriers to the integration of laboratory systems is going to take a change in mindset on the part of lab management and those working in the labs. That change will result in a significant change in the way labs work, yielding higher productivity, a better working environment, with an improvement in the return on your company's investment in your labs operations. Laboratory systems need to be designed to be effective. The points noted here are one basis for that design.

### **(Endnotes)**

<sup>1</sup> <http://theintegratedlab.com/>

<sup>2</sup> *Analytical Chemistry* is a publication of the American Chemical Society (<http://www.acs.org>) and was one of the key publications to promote laboratory automation and computing. *American Laboratory*, published by ISC (<http://www.iscpubs.com>) also was active with articles by Ray Dessy and Jonathan Titus.

<sup>3</sup> PITTCON is the largest annual trade show on instrumentation in chemistry worldwide (<http://www.pittcon.org>)

<sup>4</sup> "Are you a Laboratory Automation Engineer", J. Liscouski, *JALA* 2006;11:157–162, also available in an expanded version at <http://www.institutelabauto.org/downloads/RULAEdnld.htm>

**Institute for  
Laboratory  
Automation**  
Groton, MA  
978.449.7139

[www.InstituteLabAuto.org](http://www.InstituteLabAuto.org)  
[info@InstituteLabAuto.org](mailto:info@InstituteLabAuto.org)  
a non-profit organization

© 2009 Institute for Laboratory Automation, Groton, MA, All Rights Reserved

Page 10