Building a Smart Laboratory 2012

An introduction to the integrated lab

From the publishers of
scientific computing world

One Platform, One Goal, One Winner

The quantity of data your laboratory produces continues to grow, and so do your data management needs. You need a head start. By integrating LIMS, SDMS and ELN, STARLIMS consolidates all types of data, from all types of laboratories, into a single platform. Streamline your lab with just a single web-based informatics system to deploy and support, even across widely distributed and global organizations.

Solid, comprehensive and flexible LIMS, ELN and SDMS—the informatics solution that powers your lab. **www.starlims.com**

STARLIMS®

An Abbott Company

Offices: Argentina, Australia, Belgium, Canada, China, Colombia, France, Germany, Hong Kong, Indonesia, Ireland, India, Israel, Japan, Kuwait, Malaysia, Mexico, Netherlands, Peru, Poland, Portugal, Puerto Rico, Qatar, Romania, South Africa, Spain, Sweden, Thailand, Turkey, UAE, UK, USA, Venezuela.

Contents

Introduction – The smart laboratory

There is no specific definition of a 'smart' laboratory. The term is often used in different contexts to imply a laboratory that is designed in a way to optimise it's physical layout, or that incorporates the latest technology to control the laboratory environment, or that the laboratory is using the latest technology to manage its scientific activities. For the purposes of this publication, it is the latter definition that applies.

The progressive incorporation of information technology into all aspects of laboratory operations has resulted in fundamental changes in lab work. Prior to about 1900, most scientific innovation and development was either embedded in an industrial process, or was an outcome of academic or privately initiated research. The progressive introduction of industrial R&D laboratories heralded a new era of innovation and development with an extensive dependence on the skills, knowledge and creativity of individual scientists. The evolution has continued into the 'information age' with a growing dependence on information technology, as both an integral part of the scientific process and as a means of managing scientific information and knowledge.

"The dependence of science on technology grows relentlessly"

The dependence of science on technology grows relentlessly. From the basic application of computational power to undertake scientific calculations at unprecedented speeds, up to the current situation of extensive and sophisticated laboratory automation, black box measurement devices and multiuser information management systems, technology is causing glassware and paper notebooks to become increasingly rare in the laboratory landscape.

A frequently articulated fear about the relentless incorporation of technology in scientific processes is the extent to which it can de-humanise laboratory activities and reduce the demand for intellectual input, or indeed any fundamental knowledge about the science and technology processes that are in use.

The objective of this publication is to present a basic guide to the most common components of a 'smart' laboratory, to give some general background to the benefits they deliver and to provide some guidance on how to go about building a smart laboratory.

John Trigg
phaseFour Informatics
2012

This guide, *Building a Smart Laboratory*, is written and compiled by John Trigg, director of phaseFour Informatics, with contributions from Peter Boogaard, founder of Industrial Lab Automation, and Bob McDowall of McDowall consulting. It has been produced by Europa Science, the publishers of *Scientific Computing World,* and edited by Beth Sharp.

Unit 9 Clifton Court, Cambridge, CB1 7BN, UK.
Tel: +44 (0)1223 211 170. Fax: +44 (0)1223 213 385. Web: www.scientific-computing.com

Laboratory automation and laboratory informatics

The two primary areas of technology that apply to a smart laboratory can be broadly categorised as laboratory automation and laboratory informatics. In general, laboratory automation refers to the use of technology to streamline or substitute manual manipulation of equipment and processes, whereas laboratory informatics refers to the application of information technology to the handling of laboratory data and information.

The field of lab automation comprises many different automated lab instruments, devices, software algorithms and methodologies used to enable, expedite and increase the efficiency and effectiveness of scientific research. Laboratory informatics is the specialised application of information technology aimed at optimising lab operations. It encompasses electronic lab notebooks, sample management, data acquisition, data processing, reporting and scientific data management.

Both disciplines aim to increase productivity, improve data quality, reduce lab process cycle times and to facilitate data acquisition and processing techniques that would otherwise be impossible. Furthermore, retention and accessibility of knowledge through online storage and search algorithms aim to offer additional benefits through the re-use of existing information, the avoidance of repeating work and enhancing the ability to communicate and collaborate in real time.

The application of technology in today's labs is required to achieve timely progress and remain competitive. Laboratories devoted to activities such as high-throughput screening, combinatorial chemistry, automated clinical and analytical testing, diagnostics, large-scale biorepositories, and many others, would not exist without advancements in lab automation.

The term 'laboratory informatics' has been progressively creeping into the vocabulary of lab workers during the past decade and has come to represent the field of information technology as it is applied to a wide range of processes and operations. Typically, it addresses the convergent field of laboratory data and information systems, which includes laboratory information management systems (LIMS), electronic laboratory notebooks (ELNs), scientific data management systems (SDMS) and laboratory execution systems (LES), as well as the tools used for data acquisition and data processing.

There is a very good reason why the use of a generic term such as laboratory informatics is important; we need to get away from our traditional application-centric approach to laboratory computing and think in terms of the big picture, i.e. a fully-integrated computing environment that embraces all aspects of the application of technology to lab operations and its interaction with other company systems. This has become increasingly important as the deployment of an ELN generally represents the final step in making a lab fully electronic and hence raises the demand for interconnection between all laboratory systems. In this sense, being fully electronic and being fully integrated are two different things.

"The term 'laboratory informatics' has been progressively creeping into the vocabulary of lab workers during the past decade"

For most labs, the reality is that fully 'electronic' corresponds to an application-centric portfolio of 'systems' that were not necessarily designed to work together, and for which interoperability is hampered by the lack of standards and is therefore dependent on custom solutions. What we aspire to is an 'integrated' laboratory that is modular, based on standards and is designed to facilitate connectivity, data sharing and collaboration.

Over the past two to three years the

informatics market has experienced two interesting developments: firstly, the previously separate LIMS and ELN sub-markets have started to overlap, causing a certain amount of confusion as a consequence of our application-centric mind-set. And secondly, a number of merger and acquisition activities have reshaped the vendor line-up, specifically in the ELN field. So what do these developments mean? Do they represent something more fundamental than just functional and commercial opportunism and present some tentative steps towards addressing the integration problem?

There is no distinct boundary between laboratory automation and laboratory informatics. At one extreme, lab automation can be interpreted as a field of engineering and laboratory informatics as a field of information management, but both contribute to a common objective of enhancing the efficiency of laboratory processes.

Industry evolution and trends

During the past 40 years, the development of increasingly powerful computers has played a major role in the advancement of laboratory data and information management. Initially, the high processing capabilities of computers were exploited to perform complex calculations at unprecedented speeds, often off-line on a company's mainframe.

Gradually, as digital technologies progressed and with the development of the microprocessor, computers were brought into the lab and used for data acquisition and data processing. As a consequence, a number of laboratory techniques were revolutionised to such an extent that it is now difficult to believe, by modern standards, just how crude certain measurements had been and also what degree of confidence or accuracy they offered. The cutting out and weighing of chromatography peaks to obtain quantitative data is one such example.

The LIMS market

As computers became more prevalent in the laboratory, another of their capabilities started to be exploited, i.e. their ability to manage workflow transactions. This led enterprising scientists to develop simple, custom computerised workflow systems to operate in conjunction with data acquisition and data processing. And that basically is how laboratory information management systems were born. In the early 1980s, first generation commercial LIMS systems started to appear, usually based on minicomputers, and offered some basic functionality to support sample and test management, and reporting of results.

A second generation of commercial LIMS started to appear in the late 1980s, typically taking advantage of relational databases to provide more sophisticated functionality. The development of client-server based systems represented the next (third) generation of commercial systems, taking advantage of the evolution of the personal computer. The fourth generation emerged as the internet and wireless connectivity developed, offering opportunities to extend the reach of LIMS beyond the confines of the laboratory.

"During the past 40 years, the development of increasingly powerful computers has played a major role in the advancement of laboratory data and information management"

As LIMS products were increasingly adopted by labs, three specific additional requirements gradually became apparent. Firstly, the need to be able to transfer data from instruments directly to the LIMS in order to avoid transcription errors; secondly, the need to manage the instrument data files from which data stored in the LIMS was derived; and thirdly, the need to handle unstructured data, graphical data and collate sample data. These requirements led to the development of scientific data management systems (SDMS) and electronic laboratory notebooks (ELNs).

Functionally, the LIMS products have become increasingly sophisticated through the successive generations to the point that the dividing line between LIMS and other informatics solutions has become less clear.

The ELN market

The ELN market has been developing rapidly during the past decade, with continuing growth, but it still exhibits some degree of instability with a large number of vendors (in excess of 30 purveyors of products that purport to be an ELN) competing for market share. As a consequence, the market suffers from some degree of 'hype' (see Figure 1). Just where ELNs sit on the Gartner Hype Cycle[3] is dependent on the view you take within your organisation and your scientific domain. The general market position is probably somewhere around the 'Trough of Disillusionment', although individual vendors may occupy positions either side of this point. The 'Trough of Disillusionment' can be considered to be the turning point when we've got past the hype and can then focus on delivering true benefit. Chemistry-based and generic ELNs are probably already beyond this point, as indeed are the majority of LIMS products.

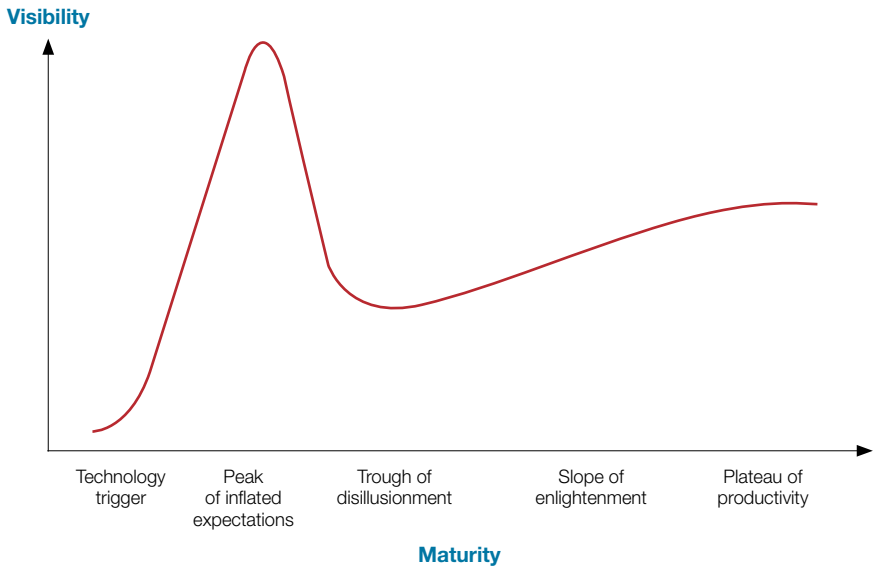
Commercial ELNs have evolved from two approaches: discipline-specific and generic. Generic software provides the architecture and tools to create and search content, and to work collaboratively in a way that satisfies the needs of almost any science-related industry. Discipline-specific ELNs are aimed at a particular market segment such as chemistry, biology or analytical. These systems are usually tailored to work with other discipline-specific software tools. Most of the commercial ELNs offer a combination of generic and discipline-specific functionality.

The initial evolution of the ELN market was centred on the provision of functionality to support small molecule chemistry. Most of the experimental processes associated with synthetic chemistry are well established, reasonably consistent and are well supported by desktop software tools. Integrating these functions in an ELN that addresses the broader capability to create, manage and store a full experimental record was a logical progression.

As a consequence, chemistry-based ELNs exhibit a good deal of maturity. If there is segmentation in this part of the market, it is determined to some extent by the origins and scope of the available products. Some, for example, will be perceived as an enterprise-wide solution, others will have more of a focus on utility and personal productivity, whilst others will provide a generic ELN capability that accepts the integration of third-party software tools.

Biology, however, has presented a bigger challenge to the ELN vendors. The more diverse and complex nature of biological processes and outcomes creates a need to capture not just the data, but also the complex interrelationships between the data. This, coupled with a diverse portfolio of biology-specific software tools, begs the

Fig. 1: The Gartner Hype Cycles



Technology Trigger: The first phase of a Hype Cycle is the 'technology trigger' or breakthrough, product launch or other event that generates significant interest.

Peak of Inflated Expectations: In the next phase, a frenzy of publicity typically generates over-enthusiasm and unrealistic expectations. There may be some successful applications of a technology, but there are typically more failures.

Trough of Disillusionment: Technologies enter the 'trough of disillusionment' because they fail to meet expectations and quickly become unfashionable. Consequently, the press usually abandons the topic and the technology.

Slope of Enlightenment: Although the press may have stopped covering the technology, some businesses continue through the 'slope of enlightenment' and experiment to understand the benefits and practical application of the technology.

Plateau of Productivity: A technology reaches the 'plateau of productivity' as the benefits of it become widely demonstrated and accepted. The technology becomes increasingly stable and evolves in second and third generations. The final height of the plateau varies according to whether the technology is broadly applicable or benefits only a niche market.

question, do biologists just need a generic ELN that will integrate with their existing software tools, or do they need a complete suite of functionality that is embedded in the ELN? The dilemma for the biologists is whether there is a commercial ELN that addresses their specific and diverse requirements. Furthermore, for those companies that need to support chemists and biologists, the question is whether it is possible to find a single vendor solution that addresses the requirements of both disciplines, or whether to choose the best of breed for each discipline.

Interestingly, within the past two or three years, there seems to be another emerging ELN domain; that of QA/QC and the regulatory world. A few vendors have concentrated specifically on this area, with products that are strongly aligned to laboratory workflows, following the step-by-step execution of SOPs (standard operating procedures) or Test Methods. The products are more structured than we would expect from a 'conventional' ELN and in some respects appear to be functionally closer to a LIMS.

This particular segment of the market has seen a number of vendors extending the functionality available in their LIMS to embrace some of the more unstructured requirements associated with experimentation. It could be argued that such products may be better labelled as 'laboratory execution systems' as they follow a very prescriptive approach applicable to those communities engaged in regulatory-based testing.

The generic functionality required just to replace a paper notebook can be a simple authoring tool capable of generating a compound-document. However, additional capability will be needed for storing and searching documents, and for addressing workflow requirements. Some organisations have chosen to implement generic ELN functionality within the framework of their standard IT tools, such as Lotus Notes and SharePoint. In the academic community, blogging tools have been used to record experimental work and thus provide the basic features of an ELN, with a strong emphasis on sharing and collaboration, and in the form of a laboratory journal.

What is laboratory automation?

Laboratory automation is defined as the use of technology to streamline or substitute manual manipulation of equipment and processes. In considering the smart laboratory, the first stage is to look at basic lab processes and computerised systems: how do they currently operate and how should they integrate? A lab may have many data systems associated with the main analytical techniques, such as chromatography, MS, UV, NIR, etc., and so can appear on the surface to be very effective, but in practice these are islands of automation in an ocean of paper.

The main way that data is transferred from system to system is via manual input, using paper as the transport medium – a slow and inefficient process. Furthermore, the process will have evolved over time and may have additional tasks that do not add any value to the laboratory output and so it becomes very slow and inefficient.

The diagnostic approach to an integrated laboratory is to map the current processes and then redesign and optimise them to use IT systems effectively and efficiently, ensuring they deliver business benefit in terms of productivity, IP protection and regulatory compliance. Therefore, the process maps for the current working practices describe what you do and why you do it. In many instances it will be due to one or more of the following:

- Custom and practice (we have always worked this way);
- Evolution over time (we have had new projects or new tasks to do);
- Extensive quality control checks (the FDA didn't like our previous way of working.

The main aim is to understand where there are bottlenecks and issues in the process – analyse and find the root causes as they will help you to challenge and improve the process. When the current process is redesigned and optimised, the aim must be

HUMAN HEALTH

ENVIRONMENTAL HEALTH

Better science happens when scientists have the best tools to work with. Ensemble® is an integrated suite of industry leading informatics solutions backed by PerkinElmer's 75 year reputation for quality and service. From R&D right through to QA/QC and manufacturing, scientists can depend on the Ensemble suite to provide the informatics tools they need to make better science happen.

Ensemble. Informatics For Better Science.

www.perkinelmer.com/informatics

PerkinElmer®
For the Better

to have, as far as it practicable, electronic ways of working and effective and efficient hand-offs and transfers between applications and organisational units.

There are three basic operating principles of the smart laboratory that should be used to redesign or optimise the laboratory processes.[1] These are:

1. Capture data at the point of origin: If you are going to work electronically, then data must be electronic from first principles. However, there are a wide range of data types that include observational data (e.g., odour, colour, size), instrument data (e.g., pH, LC, UV, NMR, etc.), and computer data (e.g., manipulation or calculation of previous data). The principle of integration must be balanced with the business reality of cost-effective interfacing – what are the data volumes and number of samples coupled with the frequency of the instrument use?
2. Eliminate transcription error checks: Never re-type data, and design simple electronic workflows to transfer data and information seamlessly between systems. This requires automatic checks to ensure that data is transferred and manipulated correctly. Where appropriate, implement security and audit trails for data integrity and only have networked systems for effective data and information sharing.
3. Know where the data will go: Design data locations before implementing any part of the laboratory automation or informatics environment. The fundamental information required is the volume of data generated by the instrumentation and where the data will be stored, i.e. in an archive system, with the individual data systems or on a networked drive? The corollary is that security of the data and backup are of paramount importance in this electronic environment. In addition, file-naming conventions are essential to ensuring that all data is uniquely identified, either manually or automatically. If required, any archive and restore processes must be designed and tested so that they are reliable and robust. The key message when designing electronic workflows is to ensure that once data is acquired it is not printed out or transcribed again, but transferred electronically between systems using validated routines. In addition, if working electronically there will be total reliance on the IT infrastructure and support systems, which need to be robust:
   - Network cabling must not have a single point of failure. Cables, switches and routers need at least two routes;
   - Sufficient network bandwidth (capacity) to handle laboratory data. Are your files 50kB CDS files or 1GB high-resolution NMR files? This needs to be designed into the network;
   - Computer hardware, especially servers and data storage devices, must be resilient and fault tolerant: dual power supplies, dual network and redundant disk storage;
   - Power backup in case of breaks in electrical supplies, not only in the computer room, but the communication cupboards for switches and routers to prevent loss of data in transit;
   - Backup and recovery systems to ensure data are not lost.

An overview of laboratory informatics

The broad relationships that exist in laboratory computing are represented in Figure 2. Although the exact terminology may vary from organisation to organisation, generally speaking, research and/or development programmes generate projects or studies. These generate experiments, which generate measurements, which in turn generate data. The triangle represents the different layers of abstraction that exist in R&D information flows. These are almost always handled by different systems.

Above the experimental layer is the 'management' context that is handled by traditional IT tools used elsewhere in the enterprise. Cross discipline collaboration tends to happen around the experiment layer. Below that is an increasing specialisation of data types and tools where only a few systems are deployed across workgroups.

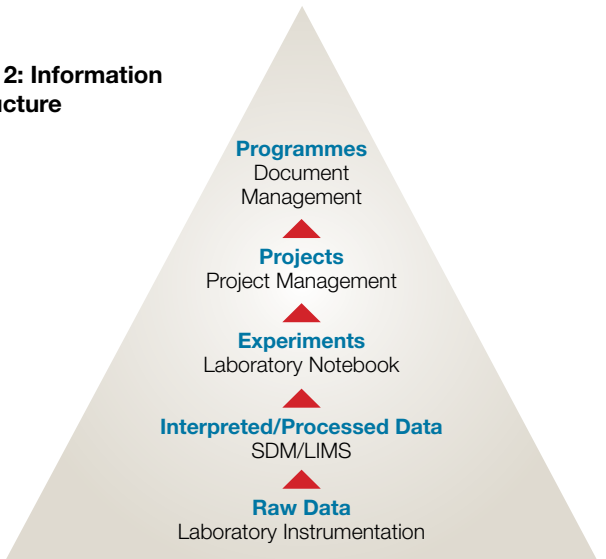
The lower three layers in the diagram represent the broad scope of laboratory informatics, embracing laboratory instrument systems, scientific data management systems (SDMS), laboratory execution systems (LES), laboratory information management systems (LIMS) and electronic laboratory notebooks (ELN).

Overall, there has been a trend towards convergence in the informatics market place. The consequence of convergence is somewhat confusing for potential customers since the term ELN is used in a very liberal sense. In fact, it is inherently ambiguous since the 'electronic notebook' is always expected to do far more than the 'paper notebook' and the additional functionality will be dependent on the type of laboratory in which it is deployed. Most of the confusion is related to the analytical and QA market segments where the differences between ELNs, LIMS and SDMS are becoming less clear. The following identifies the core differences:

ELN: Experiment-centric. An authoring tool that handles unstructured data and offers generic and specific functionality to support different scientific disciplines. Supports IP protection, knowledge re-use, productivity and collaboration.

LES: Procedure or experiment-centric. Basically able to handle structured data

Fig. 2: Information Structure



and some unstructured data. Specifically designed to meet the requirements of the GxP environment. Simplifies repeated operations. Supports electronic SOPs.

LIMS: Sample-centric. Primarily designed to handle structured data and offer sample and test management, batch operations and industry-specific workflows. Secure lab information hub. Supports compliance.

SDMS: Data-centric. Handles data files from lab instruments, meta-data, documents and the relationships between them.

Unravelling the functional and business requirements is an essential first step in any informatics project and it is easy to fall in to the trap of focusing on the solution, i.e. we need a LIMS, we need an ELN, etc., when the focus should really be on the underlying problem. Although the convergence issue in the ELN market creates some confusion, it also highlights the fact that there are a number of viable alternatives to replacing a paper lab notebook. By fully understanding the problem and identifying the functional needs of the laboratory, the solution may be found in alternative informatics applications; the challenge is to find the best overall fit for the laboratory's workflow.

One specific aspect that generally creates considerable concern during vendor assessments is the extent of the 'fit' of the vendor's product to the list of functional requirements. In some cases this may lead to a dilemma about 'buy' or 'build' and furthermore, if the decision is 'buy', what to do about the missing functionality.

As the informatics market has matured, most products offer greater degrees of configurability in order to avoid the need to write code to address any custom requirements. This enables an organisation to purchase a commercial product and configure it in ways that meet functional and cosmetic requirements, without causing any incompatibility with the core code. Customisation, involving writing code, can in some circumstances put significant barriers in the way of future product upgrades and create high costs for both maintenance and further development.

There may be instances where an organisation may choose to develop informatics solutions internally if it feels that its requirements cannot be adequately addressed by commercial products. In this case, the financial implications need to be given considerable attention. Although the development cycle, using in-house, paid-for labour, may undercut purchasing costs, it is the on-going cost of maintenance, support and the retention of knowledge that can add considerably to the total cost of ownership.

Laboratory instrument systems

The range of instruments and computerised systems used in laboratories is very wide, extending from simple instruments such as pH meters and analytical balances to sophisticated chromatographic and spectrometric systems. However, all are similar in the fact that software is used to manage the operation of the instrument, in read-only memory (ROM) or in a separate computer system to provide control of the instrument, as well as acquiring, processing, storing and reporting results.

Regardless of the type of instrument or system used, they form the foundation of the smart laboratory, as they are involved with generating the analytical data used within the laboratory. Information is abstracted from the data by a variety of mechanisms and used to make decisions. However, without the foundation layer comprising the analytical instruments and systems, the informatics portion will not function.

Simple analytical instruments

For a simple analytical instrument (e.g. a balance) the software is integral. It is capable of storing user-defined parameters or methods, but typically there is no data storage capability offered. Therefore, data is traditionally captured by either observation into a laboratory notebook or by a printout from the instrument. However, these instruments are often critical in determining the quality of the result and require a different approach to integration within a smart laboratory.

As there is no data storage capability, the instruments need to be interfaced to a system with storage functions, such as a LIMS or ELN. Interfacing needs to be direct with the receiving system, however as more intelligence is being built into simple systems, some balances have touch screens that are able to act as terminals for LIMS and ELNs.

In other cases, vendors make data capture systems to capture and buffer data from these instruments and add integrity features, such as the identity of the analyst making the measurement, data and time, sample information and audit trail entries to comply with regulatory or quality guidelines.

Computerised instrument systems

Some of the more sophisticated laboratory instruments have a separate computer system running software to control the instrument. This software is generally proprietary to the instrument vendor and has the capability to acquire, store and process analytical data and to report results. Typically, most of these instruments are designed to operate primarily in standalone mode, although there is increasing ability to transfer data to a central server for secure data storage. Examples of this type of system are chromatography data systems (CDS), UV spectrometers, mass spectrometers and nuclear magnetic resonance spectrometers.

There is typically a gap, however, in the software capability and how the system is used in a laboratory. The software of these computerised systems has the ability to process and interpret the data as well as produce the final reportable value. In practice, many labs do not use the software to its full capabilities and will print out the data and enter it manually into a spreadsheet for calculation, print out the calculation and then enter it manually into a LIMS or ELN for reporting. Obviously this is not efficient and the process needs to be redesigned to work electronically as outlined earlier: acquire data electronically at the point of origin and never retype or check for transcription error (see 'Laboratory automation').

The instrument data system must process the acquired data and then pass the reduced data or a final result to a LIMS or ELN for reporting. To ensure this is effective, some procedural controls need to be in place if the data system does not provide the functions. The primary requirement is a file-naming convention to identify files uniquely for a specific analysis or analytical run.

A potential downside of integrating systems in this way is that decision-making can be divorced from the instrument. For example, if a result is generated in a pharmaceutical lab where a computerised system is connected to a LIMS containing the product specification, the analyst knows the results is out of specification (OOS), but the data needs to be transferred to the LIMS to make the formal determination of this fact. It is a regulatory requirement that 'poor' results must not be discarded, otherwise the laboratory can be accused or testing into compliance or falsification.

Some considerations for design of computerised systems:

- Data generated by instruments must be capable of being stored directly on secure servers rather than local hard drives to protect the records;
- Databases are preferable to flat file structures as the audit trails are more encompassing and that trending is possible across analytical runs;
- Networked data systems are better than several standalone data systems of the same type. The best example of this is a networked chromatography data system where data can be acquired in one laboratory and reviewed in an office.

Most computerised systems are not designed for fully electronic working; electronic signature capability may be present to enable work to be undertaken with an electronic process, but there is little, if any, ability to hand off work electronically. As an example, if an analyst completes an analysis, would a supervisor know to review the data when they logged onto the system? In the large majority of instrument data systems the answer is no. This is a challenge for a really effective smart laboratory.

"The major business benefits of a LIMS are typically associated with increasing workflow efficiencies by eliminating manual data entry and transcription errors"

What is a LIMS?

A laboratory information management system (LIMS) provides the basic functions to address sample and test management and has become the standard tool for analytical and QC laboratories for registering samples, assigning tests, gathering and managing results, and issuing reports. With a growing level of sophistication in the relevant information technologies and configurability, most LIMS now offer a range of functionalities associated with sample and test management to provide a more integrated solution to support workflows and processes customised to a range of industry-specific requirements.

The basic functions to be found in a LIMS are the registration of samples and associated data, such as provenance, customer, due dates, etc.; the assignment of tests to the sample; scheduling and tracking of the sample and tests; recording the test procedure, equipment and materials used during testing; the review, approval and aggregation of test results for the sample; and the preparation of customer reports.

The major business benefits of a LIMS are typically associated with increasing workflow efficiencies by eliminating manual data entry and transcription errors. This is achieved through interfacing lab instruments for two-way communication of sample IDs, work lists and results, and by the integration with other lab systems such as electronic laboratory notebooks (ELN) and scientific data management systems (SDMS).

A LIMS also acts as a major repository of the records of analytical testing and can be a source of historical data associated with the organisation's products and production processes. In addition, the transactional nature of a LIMS enables a secondary record system to be maintained as an audit trail to track date, time, user and, if necessary, what change was made within the system. This data may then be used to satisfy quality assurance requirements in terms of data integrity and can also be used to generate a wide variety of management reports of the lab's performance.

Where a LIMS is used in a regulated environment, it is necessary that the system be validated and placed under change control. See 'Regulatory compliance and systems validation' for more detail on regulatory compliance.

In reality, a LIMS is more complex than just a single application, due in part to the convergence issues described in 'An overview of laboratory informatics'. In practice, the term LIMS can refer to the LIMS application; analytical instruments interfaced directly with the LIMS; lab data systems and computer systems interfaced with the LIMS (chromatography data systems, scientific data management systems, electronic laboratory notebooks, etc.); and applications outside of the laboratory that are also interfaced to the LIMS (enterprise resource planning systems).

As an example, Figure 3 shows the full scope of a computerised system that could represent a smart analytical/QA/QC lab.

Designing the LIMS environment means that you need to consider all the other systems in the lab that must interface with the LIMS. This includes other applications

Wtah's miknag yuor dtaa so hrad to depecihr?

Whether you're managing it, searching through it, tracking it, analyzing it, collaborating with it, sharing it, reporting it or protecting it – you need simple ways to get clarity and add value to your data.

You'll find the solution at **IDBS**

www.idbs.com/scw12

Enabling Science

idbs

such as scientific data management systems (SDMS), chromatography data system (CDS) and electronic laboratory notebooks (ELN), as well as various related data systems that may be interfaced to any of these systems, or that operate independently. It also includes analytical instruments, chromatographs and other sources of laboratory observations shown in the lower half of Figure 2.

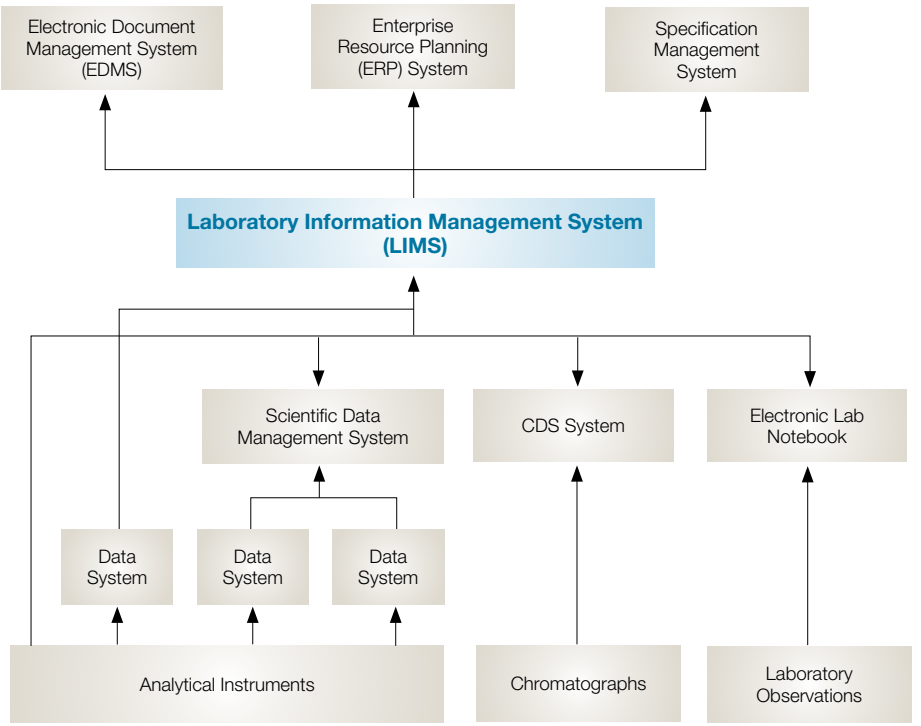
Data can be transferred to the LIMS by a variety of means:

- Direct data capture from an instrument interfaced to the LIMS;
- Data capture from an instrument with analysis and interpretation by the attached data system and only a result is transferred to the LIMS;
- As above, but the results or electronic records are transferred to the LIMS via a scientific data management system;
- Laboratory observations can be written into a notebook then entered manually into the LIMS or captured electronically via an electronic laboratory notebook (ELN) and transferred electronically to LIMS.

A pre-requisite before implementing a LIMS, or indeed any major computerised system, is to map and optimise the laboratory processes that the LIMS will automate. The lab needs to understand the process and identify any bottlenecks and their underlying cause(s).

This is especially important when moving from a paper-based to an electronic environment due to the fact that most lab processes have evolved over time rather than being specifically designed to meet organisational requirements. Moreover, the processes are paper-based rather than electronic. Refer to reference 10 for further information about the LIMS matrix. The aim for any LIMS implementation is for a simplified and streamlined electronic process, rather than the automation of an inefficient and paper-based status quo.

Fig. 3: Diagram of a LIMS environment



What is an LES?

Historically, electronic laboratory notebooks (ELN) have been designed to accommodate the unstructured data typically found in research laboratories. Researchers should be able to record scientific data, make observations, describe procedures, include images, drawings and diagrams and collaborate with others to find new chemical compounds, biological structures, etc., without any limitation. In a research environment, workflows are often methodical but unscripted, and qualitative characteristics are often most important. In QA/QC analytical labs the routine sample-processing paradigm dominates. Workflows are more repeatable and data is often much more structured and quantitative. Analysts, therefore, need a structured and robust platform to ensure that proper procedures are followed, that the progression of samples through the lab is tracked and that discrete measurement data is captured and reported reliably.

Analytical services and quality control laboratories frequently deploy systems to automate high-volume workflows and ensure compliance. In addition, many of the same needs and characteristics hold true for analysts working in core sequencing and genotyping laboratories that support R&D, clinical diagnostics labs, and the like. Traditionally in these laboratories LIMS have been very successful. In recent years, however, another category of informatics tool has emerged and is gaining significant popularity. This category of informatics products is often referred to as laboratory execution systems (LES) and they range in functionality from the simple to the quite complex.

An LES is designed to support analysts' daily workflows in a natural language form and typically provide the analyst with a User Interface (UI) that closely resembles existing lab worksheets and/or standard operating procedures (SOP). Ultimately, whether it appears as electronic laboratory worksheets augmenting LIMS functionality or as a standalone application, the LES serves the analyst by guiding them through reproducible workflows, managing the associated data and helping to ensure compliance.

An LES will extend the usability of LIMS or ERP/QM systems, because the user interface is more focused toward how the scientists operate in the laboratory. LIMS and enterprise ERP systems are often designed by technicians, business and IT specialists. Since many laboratory operations are specialised due to the scientific nature of the operation, the user experience is very critical to assure acceptance. LES are fundamentally written with the end user in mind and often operate in natural language, or mimic existing paper forms as electronic and automated equivalents.

This is often referred as 'Paper on Glass' in the market. In a QA/QC lab, standard operating procedures are developed to ensure the accuracy and consistency of sample data. It is essential that the SOP be followed exactly for each and every analysis to eliminate the risk of the measurement process being a variable. In this way, if a test result is out of range, it is more likely to be related to the sample and not the testing procedure.

LABS/Q® LIMS

Taste the Difference

iCD.

www.icd.eu * info@icd.eu

Many LES allow users to adopt their way of working on a natural way. Typical procedures include:

- Worksheet-based execution: Convert existing paper-based worksheets into electronic forms.
- SOP-based execution: Overlay data entry points, grids or other prompts on existing SOP. This supplements the SOP with real-time data collection and confirmation, in addition to ensuring that analysts are following the documented procedures.
- Define unique workflows required for routine procedures. Once the process has been mapped based on the desired workflow, the analyst can then be walked through the test process using the workflow.

Benefits of an integrated laboratory execution system (LES)
Eliminate manual data entry and reduce transcription errors
Automated SOP enforcement
Automate calculations
No user training required
Single point of truth (master record)
Supporting paperless lab initiatives

What is an SDMS?

A scientific data management system (SDMS) is, in its basic form, a system used to manage electronic records generated by laboratory instruments. Typically, an SDMS will provide facilities for long-term data preservation, accessibility and retrieval. It is complementary to other informatics systems such as LIMS and ELNs in the sense that it can provide a common repository for experiment and sample-related data files. In this way it provides a more consistent approach to managing laboratory data, dispensing with assorted local repositories and offline media, such as CDs, DVDs, tape, etc.

As with many other lab informatics tools, the lines between a LIMS, ELN and an SDMS are at times blurred through the incorporation of additional features to complement the core functionality. The basis of an SDMS is a means of collecting data files from a wide range of laboratory instruments and storing them, along with metadata, in a uniform way in a database. In other words, it is a laboratory content management system. By adding workflow elements and providing facilities for the management and storage of other documents associated with laboratory operations (worksheets, SOPs, safety information, reports, PDFs, office documents, images, etc.) in practice an SDMS can evolve into a more comprehensive single informatics solution for some labs.

However, an SDMS is essentially an 'event-driven' system that gathers data, which may limit some of its capabilities relative to the other informatics tools, and is therefore more frequently seen as a system that is complementary to a LIMS or an ELN. Nevertheless, the principle on which the SDMS is based is that it aggregates records into a logical collection associated with a specific entity such as a programme, project, experiment, product, sample, etc. to provide a readily accessible collection of relevant information. Embedded in an SDMS will also be the means to provide appropriate security of the records by means of access control, audit trail, authorisation and change management.

What is an ELN?

In its simplest form, an electronic laboratory notebook can be considered to be a direct replacement for the paper lab notebook. In this instance, it can provide the generic functionality to support 'broad' scientific documentation processes required for patent evidence creation, cross-discipline collaboration and general record keeping. However, the integration capabilities that we readily associate with information technology raise the possibility of a tighter coupling of other labsystems into the 'electronic laboratory notebook'. In other words, can the information that is currently printed from other systems, cut out and pasted into the paper lab notebook, be electronically entered or linked directly to the electronic laboratory notebook?

For example, systems that provide chemical structure drawing, structure and sub-structure searching, compound registration, etc. are an integral part of the chemistry laboratory's process and therefore would be expected to become part of an electronic solution. Similarly, other scientific disciplines will have specific requirements consistent with their particular laboratory processes. Figure 4 illustrates the relationship between 'broad' (generic) and 'deep' (specific) systems.

Another way of looking at this is to define an information structure (see Figure 2) that identifies how different systems fit into the laboratory architecture. The triangle represents the different layers of abstraction that exist in R&D information flows. These are almost always handled by different systems. Above the experimental layer is a management context that is handled by traditional IT tools used elsewhere in the enterprise. Cross-discipline collaboration tends to happen around the experiment layer. Below the experiment level there is an increasing specialisation of data types and tools, and only a few systems are comfortably deployed across workgroups.

"In its simplest form, an electronic laboratory notebook can be considered to be a direct replacement for the paper lab notebook"

From a patent perspective, the experimental layer is crucial as it captures what the scientist is thinking and doing, and therefore provides the evidence of conception and reduction to practice of the 'invention'. In broader Intellectual Property (IP) terms, it is the experiment layer that constitutes a record of the laboratory's work and as such contributes to the scientific knowledge repository. Whilst this repository resides on paper, the ability to access, collaborate and share scientific knowledge is constrained. The implementation of an ELN therefore offers a significant opportunity to bring about greater efficiencies in these processes.

The generic function of an electronic laboratory notebook supports the 'experimentation' layer and contains abstractions from the lower data levels. So in terms of 'what is an electronic laboratory notebook?', the CENSA[3] definition – 'A system to create, store, retrieve and share fully electronic records in ways that meet all legal, regulatory, technical and scientific requirements' – is all encompassing and can therefore mean different things to different people. For this reason, a clearly defined understanding of the role that the ELN is going to play in a given organisation is absolutely essential at the start of an electronic laboratory notebook project.

An ELN can serve the organisation in three ways: firstly, it can take advantage of the capabilities of IT to improve the ability to acquire, manipulate, share and store data (productivity). Secondly, it can facilitate communication and sharing in real time across multi-disciplinary and multi-site teams (collaboration). Thirdly, it can provide a scientific knowledge repository that can be easily accessed to recover records of the lab's work (content/knowledge management).

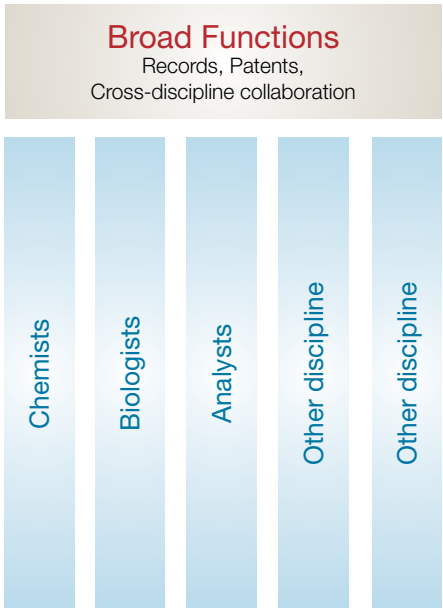
The way in which lab notebooks are used is largely dictated by the United States' patent system which, unlike the rest of the world, is based on 'First to Invent'. The need to be able to demonstrate who really was first to invent requires the laboratory notebook to be an authentic and trustworthy record that describes the concept and it's reduction to practice, and is signed by the author and corroborated by an impartial witness.

The reason why the migration away from paper lab notebooks has taken so long can most likely be attributed to two factors: the reluctance of lawyers and patent attorneys to gamble on the legal acceptance of electronic records in patent interferences and patent litigation without any case law, and the lack of confidence in our ability to preserve electronic records over several decades.

One of the more challenging barriers to a successful electronic laboratory notebook implementation is identifying exactly what role the ELN will play. The term 'electronic laboratory notebook' is inherently ambiguous. In most cases the implementation of an ELN is expected to do more than just replace the paper lab notebook. The paper lab notebook is a simple authoring tool and any electronic authoring tool capable of generating a compound document will serve as a replacement. For some companies this has proved to be the case.

The combination of Microsoft Office, SharePoint services and a means of preserving documents (e.g. in PDF – portable document format) has been shown to be an adequate replacement for paper. But if more functionality than this is needed, for example integrating various chemistry or biology-centric functions or other discipline-specific tools, then we are really talking about an electronic laboratory rather than an electronic laboratory notebook.

Fig. 4: Broad vs. Deep



Business requirements

As well as the progressive introduction of information technology in laboratories, another significant change has been the way in which the lab is managed. Business principles have largely replaced traditional scientific principles. The outcomes of laboratory work are expected to be achieved with increasing efficiency, lower costs and in shorter time frames. The combination of business-oriented science, legal compliance, regulatory compliance, corporate governance and health and safety regulations has meant that it is not unusual to find scientists feeling overwhelmed with performance targets, measures and bureaucracy rather than engaging in scientific debate.

Nevertheless, the combination of information technology and modern business practices are responsible for delivering significant increases in laboratory efficiency and productivity, yet somehow there still seems to be a level of expectation that information technology could do far more. In addition, if the technology is not right, there is a tendency for scientists to want to fix it, or find workarounds, and this can become a significant distraction from their scientific objectives.

The basic goal in deploying laboratory informatics systems is to bring about improvements in productivity and business efficiency. In order to maximise the benefits, it is important to take into consideration the wider laboratory and business processes that may be impacted by the new system. It is easy to fall into the trap of just 'computerising' an existing laboratory function, rather than looking at the broader benefits that may be accrued by re-engineering a business process. The use of tools such as Six Sigma and Lean can help considerably in this instance.

It is prudent to be careful with the use of these tools, however, depending on the nature of the lab. For example, high throughput, routine, testing laboratories, which basically follow SOPs (standard operating procedures), are more receptive to process improvement, while discovery and research laboratories, which are less structured and are dependent on more diverse and uncontrolled processes, are far less likely to benefit from formal process re-engineering.

Productivity and business efficiency are usually measured in financial terms, although this may be translated into time savings or, in some cases, numbers of tests, samples, experiments, etc., in unit time. It is necessary to be able to quote before and after figures for any deployment project, so establishing a baseline metric is an important early step in the project and can contribute significantly to the project justification.

Costs/return on investment

Any organisation considering the implementation of a new informatics or automation system will want to investigate the return on investment or cost/benefit. This is usually extremely difficult to do since most of the projected benefits will be based on a certain amount of speculation and faith. However, there are some important points to consider in building the cost/benefit case.

In the case of most lab informatics tools, the costs associated with managing paper-based processes (e.g. notebooks, worksheets, etc.) through their full lifecycle are not always fully visible or understood. Apart from the material costs and the costs of the archive process, there is a hidden cost – whilst notebooks are in the possession of laboratory scientists, for example. The time taken in writing by hand, cutting, pasting, transcription and generally manipulating paper, as well as approval and witnessing processes, all contribute to this hidden cost.

It is normal in building the cost/benefit equation to look at how much of a scientist's

time is taken up in managing the paper-based processes and to use this as a basis for potential time savings with an electronic solution. (see Figure 5). Although the start-up costs are high for an electronic solution, the incremental cost of adding new users and increasing storage space is modest.

Return on investment (ROI) tends to focus on the short term – how soon can we get a return on the money invested in deploying a new system – but the true value of the system may be long term and therefore far more difficult to measure since the value will be determined by behavioural changes. For example, there is a growing body of evidence being presented at conferences on electronic laboratory notebooks (ELN) by numerous companies that have implemented such a solution, showing that the short-term time savings are significant. In addition, each of these organisations lists a number of other non-quantifiable, long-term benefits:

- Scientists can spend more time in the laboratory;
- It is easier to find information in a searchable archive;
- It is easier to share information;
- Increased efficiency can be achieved through the elimination of paper;
- There is a reduced need to repeat experiments (knowingly or unknowingly);
- Data quality (legibility) is improved;
- A smooth transition when people leave the company;
- Online use in meetings.

If you have an ROI process, get as much help as you can to find out how it works and how to make it work successfully. Consider 'Cost vs. Value' and think carefully about the hidden costs of paper.

Regulatory compliance and systems validation

The FDA requirements for 21 CFR Part 11 compliance for most life science companies has been directed mainly in the area of development, manufacturing and clinical trials. 21CFR Part 11 outlines the criteria under which electronic records and electronic signatures are considered to be trustworthy, reliable and equivalent to paper records. Typically, research or discovery tend to be outside the direct reach of Part 11, but the principles of Part 11 compliance are closely related to legal requirements and it therefore makes good sense to take these into consideration for an ELN project.

"Any organisation considering the implementation of a new informatics or automation system will want to investigate the return on investment or cost/benefit"

In this way, the fundamental principles of the trustworthiness, reliability and authenticity of scientific evidence can be supported by a robust process that defines the course of the regularly conducted business activity used to create the evidence. Furthermore, should the ELN project be extended into development areas, then, of course, compliance will become mandatory.

In order to meet regulatory requirements, the following criteria typically apply:

- The system must be validated;
- The system should produce human readable output;
- The system should have security/access control;
- The system should have an audit trail;
- The system should have version control;
- The system should have data validity checks;
- The system should have provide an electronic signature process for all signed electronic records that includes:
  - Printed name of the signer;
  - Date and time of signature execution;
  - Meaning of signature;
- Establish corporate internal policies and guidelines for regulatory compliance:
  - Validation policy and procedures;
  - Disaster recovery;
  - Revision and change control procedures;
  - System access and security procedures;
  - Training procedures;
  - Document control procedures;
- Develop a clear, comprehensive migration strategy:
  - Include records, signatures, audit trail;
- Establish retention policies based on current predicate rule requirements.

Regulatory compliance and systems validation is more common in the GxP environment and is applied to LIMS, SDMS and LES systems.

GAMP software categories in a LIMS

The Good Automated Manufacturing Practice (GAMP) guidelines is an industry-written document for the validation of computerised systems used in the pharmaceutical industry now in its 5th version.[5] In all versions there is a classification of software into one of five categories presented in Table 1. Further discussion and debate on the GAMP software categories as applied to laboratory computerised systems can be found in the paper by McDowall.[6]

A LIMS could therefore contain the following categories of software:

- LIMS application software which is configured (category 4);
- Customisation of the product using the internal scripting language (category 5);
- Writing custom code using a recognised computer language to connect the LIMS to another application or instrument (category 5).

Fig. 5: System costs of paper notebooks and ELNs

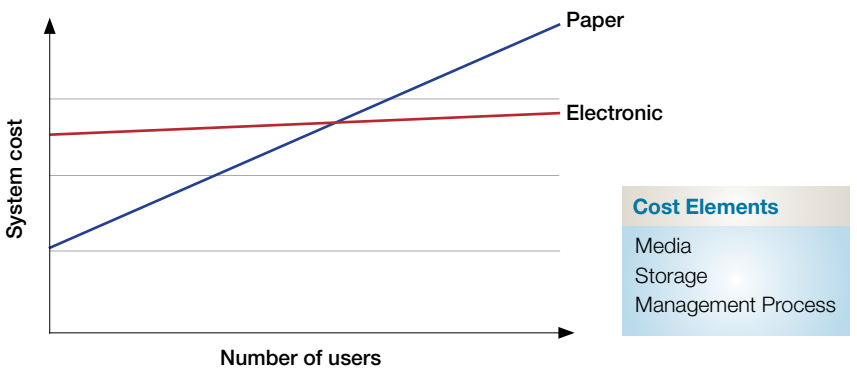


Table 1: Software categories based on GAMP 5[5]

Category 1 Infrastructure software

- Established or commercially-available layered software including operating systems, databases, office applications, etc.
- Infrastructure software tools including antivirus, network management tools, etc.

Category 2 Firmware

- Discontinued – firmware now treated as category 3, 4 or 5.
- Clash with USP <1058> over approach for Group B laboratory instruments: validate or qualify?

Category 3: Non-configured products

- Off-the-shelf products that cannot be changed to match the business processes.
- Can also include products are configurable, but only if the default configuration is used.

Category 4: Configured products

- Configured products provide standard interfaces and functions that enable configuration of the application to meet user-specific business processes.
- Configuration using a vendor supplied scripting language should be handled as custom components (category 5).

Category 5: Custom applications

- These applications are developed to meet the specific needs of the regulated company.
- Implicitly includes internal application macros, LIMS scripting language customisations, VBA spreadsheet macros
- High inherent risk with this type of software.

As a minimum, a LIMS could consist of only category 4 software, but in a GMP environment it will also contain at least one type of category 5 software, the scripting language option for customisation. This mixed environment affects the life cycle and will lengthen the time for implementation of the system. Therefore, when at all possible, the lab should change the business process to match the LIMS to reduce implementation time and validation cost, which was discussed in the previous section.

GAMP software categories and system life cycle for a LIMS

To define the risk and amount of work that we need to do when validating a LIMS, we need to understand the categories of software present. Once this is determined, the life cycle that is necessary to implement a LIMS can be defined.

System life cycle detail and documented evidence

The life cycle and documented evidence discussed in this section is based upon the validation of a number of systems, but needs to be understood in the context of an organisation's computer validation policies and procedures. Each organisation can have different approaches and terminology. Therefore, the terminology used here may be different to some organisations. However, what matters is the question of whether you performed the work described in each section rather than argued over the name of a specific document.

The key message is that you can demonstrate that the system was developed under control and is validated. The main documents needed for validation of a LIMS are presented in Table 2.

Patent-related issues

The US patent system is based on 'First to Invent' and in order to help determine who was first to invent, most companies engaged in scientific research create and preserve evidence that they can use to defend their patents at a future date. Traditionally, this evidence has been in the form of the bound paper laboratory notebook.

In a patent dispute, any inventor is assumed to have an interest in the outcome of the case, so their testimony must be corroborated. Most organisations require these notebooks to be signed by the author ('I have directed and/or performed this work and adopt it as my own') and also by an impartial witness ('I have read and understood this work').[7,8]

Evidence in US patent interferences is subject to the Federal Rules of Evidence. There are a number of important hurdles that need to be overcome, in particular the Hearsay Rule (by definition, if the author cannot be present, then the evidence is hearsay) and the Business Records Exception.

The Business Records Exception is an exception to the hearsay rule that allows business records, such as a laboratory notebook, to be admitted as evidence if they can be demonstrated to be relevant, reliable and authentic. The following criteria must be met:

- Records must be kept in the ordinary course of business (e.g. a laboratory notebook);
- The particular record at issue must be one that is regularly kept (e.g. a laboratory notebook page);
- The record must be made by or from a knowledgeable source (e.g. trained scientists);
- The record must be made contemporaneously (e.g. at the time of the experiment);
- The record must be accompanied by testimony by a custodian (e.g. company records manager).

Any doubt about the admissibility of electronic records was largely removed by this statement from the *Official Gazette* (10 March 1998)[9]:

'Admissibility of electronic records in interferences: Pursuant to 37 CFR 1.671, electronic records are admissible as evidence in interferences before the Board of Patent Appeals and Interferences to the same extent that electronic records are admissible under the Federal Rules of Evidence. The weight to be given any particular record necessarily must be determined on a case-by-case basis.'

In terms of admissibility, paper and electronic records are therefore equivalent. The judgment is made on the evidence, not the medium in which it is presented. However, it is important to understand the factors that impact upon the authenticity of electronic records and that in the adversarial nature of the courtroom, the opposing side will attempt to discredit the record, the record keeping system and the record keeping process. The integrity of the system and the process used to create and preserve records are therefore paramount.

However, a lot of organisations still

www.dotmatics.com

dotmatics
knowledge solutions

Table 2: Typical documentation for a LIMS validation

Document Name	Outline Function in Validation
System Risk Assessment	• Documents the decision to validate the LIMS or not and the extent of validation work to be undertaken
Validation Plan	• Documents the scope and boundaries of the validation effort • Defines the life cycle tasks for the system • Defines documentation for validation package • Defines roles and responsibilities of parties involved
Project Plan	• Outlines all tasks in the project • Allocates responsibilities for tasks to individuals or functional units • Several versions as progress is updated
User Requirements Specification (URS)	• Defines the functions that the LIMS will undertake • Defines the scope, boundary and interfaces of the system • Defines the scope of tests for system evaluation and qualification
System Selection Report	• Outlines the systems evaluated on paper or in-house • Summarises experience of evaluation testing • Outlines the criteria for selecting chosen system
Functional Risk Assessment and Traceability Matrix	• Prioritising system requirements: mandatory and desirable • Classifying requirements as either critical or non-critical • Tracing testable requirements to specific PQ test scripts
Vendor Audit Report	• Defines the quality of the software from supplier's perspective (certificates) • Confirms that quality procedures matches practice (audit report) • Confirms overall quality of the system before purchase
Purchase Order	• From supplier quotation, selects software and peripherals to be ordered • Delivery note used to confirm actual delivery against purchase order • Defines the initial configuration items of the LIMS
Configuration Specification	• Defining the configuration of the system policies • User types and access privileges • Default entries into the audit trail defined
Software Module Specifications	• Specifying a custom module and how it will integrate within the LIMS • Coding and documenting the module to pre-defined standards • Informal developer testing and correction of the module code
Technical Architecture (Technical Specification)	• IT platform(s) defined, e.g. terminal servers, database server together with resilience features • Operating systems and service packs • Operating environments: production, validation, etc.
Installation Qualification (IQ)	• Installation of the components of the system by the IT and the LIMS supplier after approval • Testing of individual components • Documentation of the work carried out
Operational Qualification (OQ)	• Testing of the installed system • Use of an approved supplier's protocol or test scripts • Documentation of the work carried out
LIMS Application Configuration	• Configuration of the LIMS application according to the configuration specification
Data Base Population	• Controlled input of methods to the LIMS • Controlled input of raw material, intermediates and in-process control sample and finished product specifications to the LIMS
Module Testing and Integration of Custom Software	• Formal testing of the module against the software design specification • Integration testing with the LIMS application
Data Migration	• Identification of the data elements and fields to migrate from an old LIMS, e.g. specifications, results, ongoing stability studies • Planning and executing the work • Confirming the successful data migration
User Acceptance Test (e.g. PQ) Test Plan	• Defines user testing on the system against the URS functions • Highlights features to test and those not to test • Outlines the assumptions, exclusions and limitations of approach
PQ Test Scripts	• Confirmation of software configuration • Test script written to cover key functions defined in test plan • Scripts used to collect evidence and observations as testing is carried out • Documents any changes to test procedure and if test passed or failed
User Training, SOPs and System Documentation	• Procedures defined for users and system administrators including definition and validation of custom calculations, input of specifications, account management and logical security • Procedures written for IT related functions • Practice must match the procedure
Service Level Agreement (SLA)	• Agreement between the laboratory and IT for IT and infrastructure services for the LIMS
User Training Material	• Initial material used to train super users and all users available • Refresher or advanced training documented • Training records updated accordingly
Validation Summary Report	• Summarises the whole life cycle of the LIMS • Discusses any deviations from validation plan and quality issues found • Management authorisation to use the system • Release of the system for operational use (this can be a separate release certificate in some organisations)

require their scientists to keep bound laboratory notebooks. This is due to the fact that there isn't the case law and/or other experience for most legal advisors to feel as comfortable with electronic records as they are with paper. The issue is not one of admissibility, but of the weight that the record will have in court. Unfortunately, we are unlikely to see a suitable body of case law for a great many years.

The high-stakes nature of the problem, lack of experience and long-term accessibility concerns have caused a number of organisations to adopt a hybrid solution – using an ELN front-end tool to author records and then preserving the resulting records on paper. This gives the benefits of paper records (for the lawyers) whilst providing the scientists with the benefit of new tools. A fully electronic system will require scientists to sign documents electronically and for the resulting record to then be preserved electronically.

Using multiple systems for patent evidence creation and preservation can expose an organisation to increased risk due to the need to maintain the integrity of each system and the consistency of the content between them. Similarly, the use of generic systems for such a task can increase discovery concerns as well as the likelihood of problems occurring. Further guidance should be sought from records management personnel and legal advisors within the organisation in order to determine policy.

A recommended approach to help uncover and resolve legal/patent concerns is to work with lawyers and patent attorneys to simulate the presentation of ELN evidence in the court room and work back to the creation of that evidence in the laboratory.

The America Invents Act – implications

Patent reform legislation, in the form of the Leahy Smith America Invents Act 2011, will change the US system from First to Invent to First to File in March 2013. It is very tempting to view this change as an opportunity to relax some of the procedural requirements of ELNs used in research laboratories. However, there are clauses in the Act that would suggest it is wise not to make such an assumption. It is likely that patent interferences and interfering patent actions will continue for many years for patents and applications filed after March 2013.[10]

There are specific circumstances described in the America Invents Act that, for example, require proof of inventive activities to remove prior art for joint research activities, or preserve the right to an interference if the application contains, or contained at any time, a claim to an invention filed before March 2013. Until the act becomes effective, and there is clarification about the implications of the new legislation, there is no reason to either change in-house procedures for keeping laboratory notebooks, or any reasons for vendors to revise the procedures and workflows in their ELN products. The more immediate concerns are:

- There is a loophole that will allow people to prosecute a patent under the old First to Invent rules for many years to come. First to File isn't dead after March 2013;

- There are some changes that mean proof of inventive activities will be especially important for joint research activities. You may also need to improve the retention of other documentation related to Joint Research project;
- Derivation proceedings will require proof of inventorship.

To add further uncertainty, there's always a chance – or indeed probability – that things are going to end up in the Supreme Court to examine the constitutional implications of a move away from First to Invent. So it does appear that the new Act makes legally robust, signed and witnessed records of inventive activities (generally in the form of lab notebooks) even more critical.
With a move to First to File, there's the additional pressure of getting to the Patent Office quickly, which means that it is necessary to start paying attention to their patent filing process, which has historically not been under much time pressure.

Data integrity, authenticity and management

Whenever electronic records are used within the framework of legal or regulatory compliance, data integrity and data authenticity are fundamental requirements of the computer systems used to create, manipulate, store and transmit those records. These requirements may also apply to in-house intellectual property (IP) protection. It will therefore be necessary for a laboratory informatics implementation project to very carefully consider the specific requirements of their organisation in this area.[11]

Data integrity

Data integrity, in a general sense, means that data cannot be created, changed or deleted without authorisation. Put simply, data integrity is the assurance that data is consistent, correct and accessible.

Data integrity can be compromised in a number of ways:

- Human error during data entry;
- Errors that occur when data is transmitted from one system to another;
- Software bugs or viruses;
- Hardware malfunctions;
- Natural disasters.

There are many ways to minimise these threats to data integrity. These include:

- Backing up data regularly;
- Controlling access to data via security mechanisms;
- Designing user interfaces that prevent the input of invalid data;
- Using error detection and correction software when transmitting data.

Data authenticity

Data authenticity is the term used to reinforce the integrity of electronic data by authenticating authorship by means of electronic signatures and time stamping. An electronic signature is a generic term used to indicate 'an electronic sound, symbol or process, attached to or logically associated with a record and executed or adopted by a person with the intent to sign the record.' Generally speaking, they are considered admissible in evidence to assure the integrity and authenticity of electronic records.

A digital signature is a specific sub-set of an electronic signature that uses a cryptographic technique to confirm the identity of the author, based on a user name and password and the time at which the record was signed.

The requirements for an informatics project will be somewhat dependent on the nature of the organisation's business and internal requirements, but security, access control and electronic signatures are factors that must be given appropriate consideration.

Data management

There are a number of ways that we can assure data integrity and authenticity. The first is to develop clear written policies and procedures of what is expected when work is carried out in any laboratory – the integrity of the data generated in the lab is paramount and must not be compromised. This is the 'quality' aspect of the quality management system (QMS) that you work under. There is the parallel need to provide initial and ongoing training in this area. The training should start when somebody new joins the laboratory and should continue as part of the individual's ongoing training over the course of their career with the lab.

To help training staff we need to know the basics of laboratory data integrity and the main criteria are listed below:

- **Attributable:** who acquired the data or performed an action and when?
- **Legible:** can you read the data and any laboratory notebook entries?
- **Contemporaneous:** documented at the time of the activity;
- **Original:** written printout or observation or a certified copy thereof;
- **Accurate:** no errors or editing without documented amendments;
- **Complete:** all data including any repeat or reanalysis performed on the sample;
- **Consistent:** all elements of the chromatographic analysis, such as the sequence of events, follow on and are date- or time-stamped in the expected sequence;
- **Enduring:** not recorded on the back of envelopes, cigarette packets, Post-It notes or the sleeves of a lab coat, but in laboratory notebooks and/or electronically by the chromatography data system and LIMS used in the lab;
- **Available:** for review and audit or inspection over the lifetime of the record. Laboratory staff need to understand these criteria and apply them in their

respective analytical methods, regardless if working on paper, hybrid systems or fully electronic systems.

To support the human work, we should also provide automation in the form of integrated laboratory instrumentation with data handling systems and laboratory information management systems (LIMS) as necessary to perform the work. In any laboratory, this integration needs to include effective audit trails to help maintain data integrity and monitor changes to data. Supervisors and quality personnel must monitor these audit trails to assess the quality of data being produced – if necessary a key performance indicator (KPI) or measurable metric could be produced.

Knowledge management

When it comes to purchasing and implementing laboratory systems software, return on investment is inevitably one of the key drivers. The up-front requirements to justify the expenditure are usually aligned to process improvement and productivity. However, there are often secondary and unquantifiable requirements about improving knowledge management in the organisation by sharing and making lab information accessible across departments, sites and geographies.

'Knowledge management' is the term being applied to the processes that address this gap. The term carries with it some semblance of a business fad since it comes with its own language and mystique, but the reality of business dependence on knowledge means that there is a growing requirement to do something to manage the organisation's intellectual capital.

Knowledge management is a business initiative that has become increasingly familiar in recent years amongst organisations striving to compete effectively in the information age. In some respects knowledge management is a logical progression from TQM (total quality management) and BPR (business process reengineering). TQM defines the concept of continuous incremental improvement (doing things well) through a data-driven, statistical quality control approach. BPR represents a paradigm shift (doing things better) by introducing high level, re-designed, integrated processes that exploit information technology.

Knowledge management extends these processes by addressing the contribution that the organisation's collective knowledge can make (doing better things) by taking into account the skills, expertise and personal knowledge of the workforce. It is perceived by many as a mechanism for manufacturing and service operations to bring about business transformation through alignment with the benefits and demands of the 'information age'.

What is knowledge management? Very simply, it is a term used to describe the processes which bring people and information together to address the acquisition, processing, storage, use and re-use of knowledge to develop understanding and to create value. The ultimate objective is to improve the performance of the business

"The ultimate objective is to improve the performance of the business and maximise the value of its intellectual capital"

and maximise the value of its intellectual capital. In some respects, knowledge management is part of a continuum that starts with data management and progresses through information management. However, the transition between data and information is governed by rules and context.

The transition between information and knowledge is governed by context, but also by the application of a number of human qualities such as insight, understanding, intuition, skill and experience. It is this human element that sets knowledge management aside from TQM and BPR; knowledge management is about people, not about information technology.

Definitions, theories and strategies about knowledge management abound to the point of confusion. It comes with, inevitably, its own taxonomy and semantics to add to the confusion. However, the basis for knowledge management can be represented by three fundamental components:

1. Enabling technologies (typically information technology – but it could be pencil and paper!);
2. People (the organisation, its behaviours and culture);
3. Processes which bring people and information together (the knowledge processes).

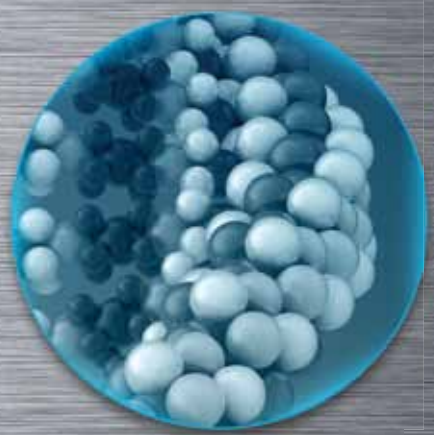
The principles of knowledge management (KM) all make good sense, it's just that an 'industry' seems to have grown around the topic that sees it as a potential revenue stream. The following conclusions add some perspective:

- KM solutions do not come in a shrink wrap box;
- You cannot implement KM, it is an outcome;
- KM is about people – technology can facilitate good KM, but that's all. Basically, information technology is a big part of the problem, but a small part of the solution.

Longer-term benefits accrue from sharing and making information accessible, ensuring that systems are easy to use, and evolving a culture based on collaboration. In terms of the smart laboratory, three simple rules are:

1. Align knowledge management initiatives with business strategy – KM has, eventually, to deliver some bottom-line results otherwise it will have no credibility;
2. Integrate knowledge management processes into the corporate culture – if it's not 'the way we do things around here', in terms of behaviour and culture, it will not get any sustainable buy-in;
3. Deliver the right information to the right people at the right time – if the 'technology enablers' are not doing the right things, the strategy will collapse.

POWERING INNOVATIONS IN CHEM & BIO



Attending Pittcon, the world's largest annual conference and exposition for laboratory science, gives you the power to get a hands-on look at the newest equipment from nearly 900 exhibitors. Participate in more than 2,000 technical presentations to hear the latest trends from industry experts and learn about recent scientific discoveries from world-renowned researchers.

For more information on technical sessions, exhibitors and short courses, **visit www.pittcon.org.**

follow us

PITTCON™
CONFERENCE & EXPO
2013 PHILADELPHIA
MARCH 17-21

Pittcon 2013 App Coming Soon

Scan this or go to www.pittcon.org to save 50% on registration.

Technical issues

Information and Communication Technology (ICT) democratises innovation and enables smaller companies, academia and students to get access to computing that was previously only available to large organisations. However, adopting new computer and infrastructure technologies in labs only makes sense if it results in better and more productive scientific operations.

The rapidly changing field of ICT can be overwhelming. Laboratory work is becoming increasingly collaborative and complex, leveraging multiple technologies to improve scientific measurement techniques and scientific understanding. This exponential rise in the scale of data being generated, combined with the increased collaboration, has resulted in a need to rethink how data is stored, analysed and shared cost-effectively.

The primary requirement for the deployment of lab informatics systems is that they can be integrated within the organisation's existing IT infrastructure. The IT department will be able to advise how this can be achieved and how the proposed solution will integrate with existing investments. There are a number of factors to be taken into consideration in the broader set of requirements.

It will be important to take geographical distribution into consideration. In wide area networks, latency becomes the most noticeable issue and acceptable maintenance windows may disappear if the implementation extends over multiple time zones. In time, the system will need to accommodate growing volumes of data, both in terms of the number of data items and the size of individual data items. Additionally, the number of users will generate more disparate requirements, not the least of which is the concern of how many people will get upset if the system goes down.

Systems architecture

Multi-user informatics systems are typically based on two- or three-tiered structures in which the application software and database may share a server, or be located on separate servers, with the client-side software deployed on a local desktop, laptop or mobile device. Traditionally, the servers are based in-house, but hosted services (cloud/SaaS) are generating increasing interest, based on potential business benefits.

From the user perspective, the client-side options fall into the following categories: thick-client and thin-client. The thick client is usually a substantial software installation on a local computer in which a good deal of the data processing is undertaken before passing the output to the database server. This has the advantage of distributing the total processing load over a number of clients, rather than the server, and may also allow a certain amount of personalisation of the client software to support individual users' needs.

The downside is that system upgrades can become time consuming and potentially troublesome depending on the local configuration, although centrally-managed systems are now making thick client systems easier to deploy, maintain and support.

Thin clients typically access the application and database server(s) through a browser. No local processing power is used, so the server and network performance are critical factors in providing good performance. The use of a browser eliminates deployment and upgrade costs, but may restrict or limit user configurability.

With regard to devices, successful deployments have been made with small form-factor PCs on the laboratory bench, remote desktop, Citrix and a KVM switch operating between a desk-bound processor unit with keyboards and screens on the desk and in the laboratory.

Mobile devices, such as tablets, laptops and smartphones, have been less successful

in fixed lab environments. However, they offer an effective solution where there is a specific need for mobility. The increased market penetration of smartphones and tablets in the consumer market has led to a rising interest in their potential in a business context through the deployment of dedicated apps or browser-based access. In general, small format handheld devices offer good access to data and information, but are typically limited when it comes to data input.

Cloud

Concurrent with this interest in mobile devices is the deployment of cloud-based infrastructures that provide hosted services. This approach brings with it opportunities to deploy rapidly at low cost with little or no capital expenditure, but does raise some questions about security, data integrity and data ownership.

Some informatics vendors already offer this type of service. Cloud services generally fall into one of two categories: public clouds and private clouds. Public clouds utilise a single code base for the service to multiple clients that limits customisation and integration, but helps keep costs down. A private cloud will typically offer a code base specific to an individual client and will accommodate customisation and integration, but will normally come at a higher management cost.

Interest and uptake of cloud solutions for informatics systems is somewhat constrained by IP, legal, regulatory and security concerns. Additionally, at the time of writing, cloud computing is suffering considerable hype – there is little doubt that in some markets the benefits of the cloud outweigh the risks, but in established laboratory environments, the case is unproven.

SaaS vs. cloud	
Cloud: All about the data	• Service level agreement • Emergency and escalation plan • Back-up services • On-demand scalability • On-demand capacity
SaaS: All about what you can do with the data	• Subscription (Pay as you go) • Free and paid services • Application consultancy offerings • Application support • Automatic updates • No investment in software • Minimal hardware investment

SaaS and cloud candidates
Prototyping which requires lots of IT infrastructure
Mobile applications
Non-core applications
Intensive collaboration projects
When scalability results in high down time
Applications requiring high pre-project investments
Projects not requiring strict 21CFR compliance
Big data projects

Cloud computing is a combination of technologies and service offerings that has the potential to increase the speed of basic research projects significantly. The effort to build such a high-performance computing infrastructure has been significantly changed from weeks to minutes and doesn't require traditional IT staffing anymore, since it is all offered as an external service.

Pay by the cycle

The cloud as an infrastructure gives researchers computational access on a subscription or pay-by-demand cost structure. Instant access to computational power, significant lower administration costs, no capital spending headaches and no dependence of availability of IT resources. Software as a service (SaaS) is a software delivery model in which software and its associated data are hosted centrally in the cloud and are typically accessed using a thin client computer or tablet device, or using a web browser over the internet.

Zero footprint applications imply that no software needs to be pre-installed on your client. This significantly helps to simplify installation procedures – a browser is all you need. To successfully upload large datasets, it is critical to have access to fast network infrastructures. Limited network bandwidth, especially in start-up phase, may result into frustration and should be avoided. As a researcher, you want to avoid thinking in computer terms.

The table on page 27 summarises the overall nomenclature and major acronyms for the most common service models.

Operational vs. investment budgeting

From a financial perspective, the model is attractive. Instead of spending significant capital and subsidising a lengthy budget and planning process, a system can now be funded on an expenses budget, using a pay-as-you-go pricing model, which can mean that you stop paying for equipment that sits idle between experiments. The cloud will significantly increase the speed of executing big data computing in research because there will be almost no wait for approving costly IT budgets. There is also no need to invest in perpetual and expensive software licenses. The cloud has the potential to be the best new development going mainstream for scientific researchers – access to an almost unlimited amount of computer power to achieve scientific calculations and text searches, combined with an almost unlimited size of disk space for an affordable price.

Data storage

Most of the short-term benefits of deploying an informatics system are associated with personal and laboratory productivity; however, the long-term benefits may accrue from the accumulated content of the system's information/knowledge repositories. This may raise further IT considerations over time with regard to how this information is managed and used. This means that the provision of adequate data storage space must be taken into account. As the volumes of data grow, there is likely to be an increasing need for better search and visualisation technologies than are typically available today.

In addition, consideration must be given to the nature of the data and how it can be efficiently stored, retrieved and interpreted. For these purposes, it is necessary to distinguish between the content of the ELN (experimental write-up) and external data (laboratory data) to which the write-up may be electronically linked. Over time, they may present two separate data preservation problems. Lab data is often stored in proprietary formats, so forward compatibility to future application and operating system releases will be critical. Additionally, the location of these files will need to be managed carefully to avoid breaking electronic links.

Electronic records management is

Software as a service licensing types	
SaaS	Software as a service: Software distribution model in which discrete applications are hosted by a vendor or service provider and made available to customers over the internet.
PaaS/STaaS	Platform/Storage as a service: External service that provides the hardware, operating system, software upgrades, security and everything else related to the day-to-day hosting of an (enterprise) application.
IaaS	Infrastructure as a service: Provision model to outsource equipment used to support operations, including storage, hardware, servers and networking components.
DaaS	Desktop as a service: The outsourcing of a virtual desktop infrastructure to a third-party service provider.
XaaS	Anything as a service: Refers to an increasing number of services that are delivered over the internet rather than provided locally or on-site.

therefore a key function to be considered from the start of a project and will require the assignment of adequate resources, from an IT perspective as well as a records management perspective in order to realise the long-term benefits of a growing knowledge repository (see 'Electronic records management').

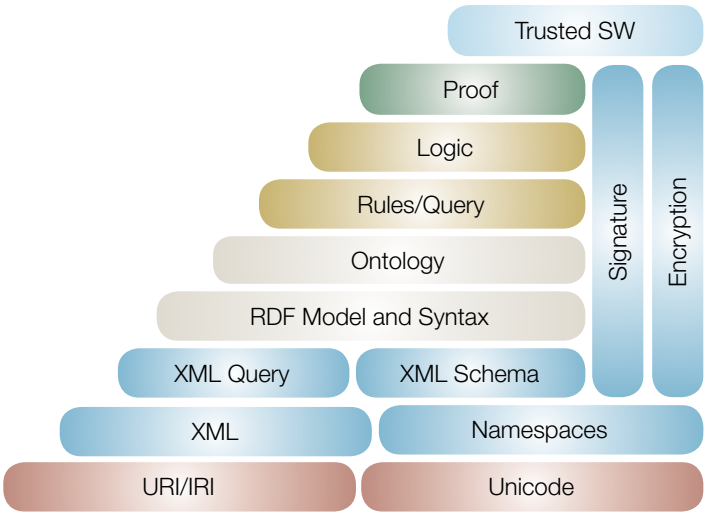
Data integration

The introduction of an ELN to a laboratory raises the question of how easy is it to make data available to the ELN from other lab systems. For example, do the data formats need to be converted to make them accessible? How easy it is to get results, spectra, chromatograms, images, reports, etc. stored on a LIMS into the record of an experiment stored in an ELN, and how easy was it to get the results, spectra, chromatograms, images, etc. into the LIMS in the first place? The question of integration is paramount to the future of efficient lab information management, not only from the perspective of record keeping, but also for productivity and real-time collaboration. The issue here is not the format of the final document, but the format of the underlying data and this is an area where proprietary file formats have dominated.

If we step outside the lab for a moment and look at some of the developments in other domains, one of the most striking advances is occurring in the framework of collaboration and the role of the internet. The use of social networking tools has

"Cloud computing is a combination of technologies and service offerings that has the potential to increase the speed of basic research projects significantly"

Fig. 6: Semantic computing



illustrated how technology can enhance collaboration once some rudimentary data standards are in place. The relative ease of incorporating text, audio, images and video into a compound document (web page) with links to other documents and references to similar documents is, in principle, akin to the role of the ELN. Furthermore, this approach does not require much more than an internet browser to gain access to the document, which means that the document is easily available from anywhere with internet access and is device-independent; i.e. any device that can run a browser should, within reason, be able to access, display and provide editing capability to the document.

The internet and web-based tools

The continuing evolution of the internet is having a significant effect on laboratory informatics with an increased need to deploy web-based tools to support geographically dispersed communities and mobile users. Web 2.0 was the generic term used to identify the change in the role of the internet as it became a collaboration space rather than just a presentation space. There were two basic principles behind this. The first was that software applications would run on the web itself rather than on the desktop. This changed the nature of the web from a collection of destination sites to a set of sources, both of data and of functionality, that could all be integrated to create the applications we need. The second principle was that of participation; of people making the web applications more useful even as they worked with them. This usually meant that they contributed their own knowledge by sharing the data as they manipulated their way through these applications.

The semantic web promises even more in the sense that adding metadata, i.e. more meaning, to the data and information on the web will enhance the capability of the technology to understand content and therefore take some of the burden of intelligently finding and filtering information within a given context. If this goal is successful, it promises further personal productivity gains.

Four laws of computing – capacity and requirements	
Moore's Law (1965)	Formulated by Gordon Moore: 'Every 18 months to two years, twice as many transistors can be fitted onto a chip of any given area for the same price' (computers become faster and the price of a given level of computing power halves every 18 months)
Metcalfe's Law (1980)	Attributed to Robert Metcalfe: 'The value of a network is proportional to the square of the number of nodes' (as a network grows, the value of being connected to it grows exponentially, while the cost per user remains the same or even reduces)
Gilder's Law (1990)	Proposed by George Gilder: 'Bandwidth grows at least three times faster than computer power' (if computer power doubles every 18 months, then communications power doubles every six months)
Zuckerberg's Law (2011)	Every year, for the foreseeable future, the amount of information you share on the web will double

Semantic computing

Research is an extremely information-intensive endeavour. As a rule, experimental data is both dynamic and distributed. New data and properties of research samples and models are continuously invented. It becomes very difficult to track the entire experiment from hypothesis and methods to the raw data, to the processed and analysed data, through to the results, conclusions and reports, and then to the management decision processes to find and use the information. As data becomes distributed into cloud repositories, it becomes increasingly important to use semantic methods to annotate and be able to relate these data to other scientific information.

Ontologies and hierarchically-organised controlled vocabularies are fundamental to helping the subject matter experts find the right information (Figure 6). Cloud computing and semantic web technologies are now becoming mainstream in leading research organisations and we have been applying these during the past decade to turn data into shared and actionable knowledge. The impact of semantic web technology in the cloud is significant and will considerably decrease the need for central warehouses of information. There will be far less need to transfer large amount of data across networks.

Electronic records

One of the most fundamental issues in replacing paper with electronic media is the concern over long-term data preservation.

This is an issue that every electronic records initiative has to face, regardless of scope or scale. Whether it is government records, medical records, scientific records or a personal collection of digital photographs, the problem is the same; what do we have to do to ensure that we can still access these records in 50 years' time, or more? The track record of the IT industry in addressing long-term data preservation is not good; proprietary data formats, content tied to the application, lack of integration standards, unreliable media and a lack of a stable operating system environment all conspire to produce a major challenge in switching to electronic records.

For laboratory informatics, there are two major factors to take into consideration to guarantee the long-term preservation of records: data formats and electronic records management.

Data formats

Currently the most common solution for electronic laboratory notebooks (ELNs) is to render completed experiment documents as PDF or PDF/A files for long-term record keeping. The Portable Document Format (PDF), created by Adobe Systems, has evolved into an international standard (ISO 32000) for rendering documents in a fixed layout that is device-independent. PDF/A is a standard for the long-term archiving of electronic documents and is also an ISO standard (ISO 19005-1:2005). It is a subset of PDF, leaving out features not suited to long-term archiving.

The use of PDF or PDF/A is generally considered to be the best option for the preservation of documents for IP and legal purposes, but the long-term storage of lab data is a different matter. The biggest problem by far is proprietary data formats. Most laboratory instruments are not designed with long-term record keeping considerations in mind and generate data in proprietary formats. In some markets, this situation is being progressively addressed by means of open platforms that provide data interchange and systems' integration standards. In the clinical world there is good progress, driven by the demand of health records. Unfortunately, the laboratory world does not have an equivalent driver.

For this reason, it becomes important to establish not only how the laboratory records are managed, but also how data files are stored and where. Although most commercial ELNs will accommodate data files in the ELN database, best practice is generally to store data in a separate repository, possibly an SDMS, with appropriate links, and to include appropriate graphical representation of the data in the ELN; basically the electronic equivalent of pasting in selected graphical instrument output into a paper lab notebook.

In this way, the ELN contains sufficient information to support the conclusions of the experiment and to allow the experiment to be repeated, but eliminates the need to manage disparate proprietary data within the ELN system. Where data resides outside of the ELN, links to this data can be made from within the ELN, but it will be necessary to take into consideration how this data and the links are managed over the long term through new revisions of software.

Electronic records management

Long-term data preservation raises the question whether the records should be stored in the ELN, or whether they are transmitted to a separate records management system, leaving the ELN to serve mainly as an operational system. For large companies, records management is typically a broader organisational process, managed by specialists, and ELNs often need to be integrated with an existing records management system. For smaller companies, it will be an essential decision in the implementation project to determine just how ELN records will be managed over the long-term.

Basic records management guidelines are as follows:

- Integrate electronic content/records management;
- Understand the legal implications of electronic records;
- Establish a file plan;
- Establish an electronic records preservation file plan;
- Establish a records management team;
- Train the technical team;
- Establish and communicate policies;
- Avoid point solutions;
- Don't keep electronic records forever.

General requirements for electronic records are summarised in the table below.

Systems integration

Prior to around 1900, most scientific innovation and development was either embedded in an industrial process or was an outcome of academic or privately-initiated research. The progressive introduction of industrial research and development labs heralded a new era of innovation and development with an extensive dependence on the skills, knowledge and creativity of individual scientists. The evolution has continued into the 'information age' with a growing dependence on information technology as both an integral part of the scientific process and as a means of managing scientific information and knowledge.

From the basic application of computational power to undertake scientific calculations at unprecedented speeds, to the current situation of extensive and sophisticated laboratory automation, black box measurement devices and multiuser information management systems, technology is causing glassware and paper notebooks to

	Quality criteria	Material to be preserved	Scrutiny	Timescale	Format
Internal use	Internal requirements	Scientific data and experimental write up	Little, if any	Company defined	XML, De Facto standards, ad hoc
Regulatory 21CFR part 11	Published regulations, with comment	Primarily scientific data. Some write up	Regulatory inspection	Defined by regulations	XML, De Facto standards
Patents	Case law, Federal Rules of evidence, etc.	Primarily experimental write up. Some data	Adversarial	~ 10 years before first come under scrutiny. Retain for 50 to100 years	PDF (paper, microfilm)

become increasingly rare in the lab landscape.

However, the progressive introduction of information technology into laboratories has been conducted in a piecemeal fashion, generally leading to the adoption of best-of-breed solutions for specific lab requirements. The consequence is that laboratories have a plethora of systems that were not necessarily designed to work together, that are based on proprietary platforms and that adopt proprietary data formats.

The deployment of an electronic laboratory notebook has been the catalyst for an increasing demand for systems integration. Up to now, integration has largely been achieved by bolting together various disparate systems that meet specific functional requirements and which often end up as a group of interconnected silos. Although this has been an important stepping-stone that has often delivered significant productivity benefits, it comes at the cost of significant effort and with the worrying legacy of the on-going management of a custom solution.

This extensive adoption of technology presents a difficult ROI calculation due to the growing need to transfer and share data seamlessly between systems. Despite all of the outstanding advances, it would be quite revealing to add up the total cost to industry to create and maintain custom solutions and middleware to solve integration problems, in addition to how much time has been wasted in not having direct and immediate access to data that is locked in inaccessible systems and how many risks have been taken in using all kinds of crude and insecure methods of transferring data. To some extent this is a legacy issue as most of our current systems were not necessarily designed to work together. In addition, systems often do a poor job of separating the content from the functionality, thus making the integration challenge even more difficult.

The demand for integration has been a technology driver behind a number of merger and acquisition activities in the informatics market during the past couple of years. This may well lead to progress, although it will be tempered by the fact that scientists usually opt for best-of-breed products, rather than single vendor solutions.

The lack of integration standards is a well-established problem in the laboratory world, and there's very little evidence of a universal solution emerging. Ongoing efforts with AnIML[12] and the Pistoia Alliance[13] have yet to gain enough inertia to make a realistic difference. SiLA (Standardization in Lab Automation)[14] is attempting to introduce new interfaces and data management standards to facilitate the integration of lab automation systems. The IQ Consortium (International Consortium for Innovation and Quality in Pharmaceutical Development) is an international association of pharmaceutical and biotechnology companies aiming to advance innovation and quality in the development of pharmaceutical products through scientifically-driven best practices and standards.[15] Furthermore, the Institute for Laboratory Automation[16] is proposing a collaborative programme of activity to drive the development and adoption of data interchange standards.

Functional/user requirements

Gathering user or functional requirements is one of the key tasks, usually assigned to the project team, for providing a specification against which potential solutions can be evaluated. The task involves uncovering and understanding user needs, distinguishing them from 'wants' and 'nice to haves', and aggregating the needs into a requirements specification.

In this context, reference to 'users' includes not just end users of the proposed system, but anyone who will interact with the system, or be involved with inputs or outputs to the system. In order to do this, various methods may be used to gather needs and to prioritise them.

In principle, the requirements specification should define what the solution must do both in terms of its functionality and how it should perform in the customer's environment.

The requirements may include, but are not limited to:

- General business requirements;
- User/functional requirements;
- IT requirements;
- Interface requirements;
- Regulatory issues;
- Data management requirements;
- Error handling;
- Reporting requirements;
- Performance requirements.

The criteria that define required performance may include:

- Access control and security;
- Look and feel;
- Robustness;
- Scalability;
- Ease of use;
- Technical performance/ response times;
- Technical support.

All of these requirements are normally collated into a request for proposal (RFP) that will be submitted to potential vendors. The RFP should also provide more general information, including an introductory description of the organisation and the major objectives of the project and diagrams showing relevant workflows. It may be preceded by a request for information (RFI), a means of gathering information about a potential vendor's products and services, which may be used to fine tune a final list of vendors to whom the RFP may be submitted.

Unfortunately, users are notoriously bad at telling what they need. Most systems are specified or designed by a team or committee and the team/committee members tend to be volunteers who are committed to the concept of the system, enthused about the improvements it can bring and can envision the potential. Unfortunately, the committee process can create complex systems, reflect compromises

"Gathering user or functional requirements is one of the key tasks for providing a specification against which potential solutions can be evaluated"

"The demand for integration has been a technology driver behind a number of merger and acquisition activities in the informatics market during the past couple of years"

It seems that one of the unfortunate consequences of the lack of industry standards is that there is the risk that too many disparate efforts to solve the problem may leave us with too many standards. However, another way in which we may find the integration issues to be solved is through the evolution of generic technology standards. Advances in consumer information technologies have a strong focus on communication, sharing and collaboration – three criteria that are essential to integration in laboratories. In the consumer world, the internet is the platform and the World Wide Web Consortium manages the standards that facilitate satisfying these criteria. The progressive adoption of web technologies for laboratory informatics may well enhance the possibilities of achieving greater levels of integration.

and it is often the case that most problems come from people who don't volunteer for the committee.

By definition, the members of the team are more committed to the success of the project than those who are not directly involved. In their deliberations, project teams often develop a concept of a solution which is much more sophisticated than might be needed or indeed is economically justifiable.

Typically the 'requirements gathering' phase involves harvesting needs, wants and ideas from the potential user community and then engaging in a prioritisation exercise to reduce the list to a specific set of requirements that form the basis of a request for proposal (RPF) that can then be presented to vendors. It is important that the business requirements are fully clarified; this ensures that the scope of the project is defined and can therefore help exclude some of the more exotic 'needs' that might arise. Any single item on the requirements list should justify itself not only financially, but also in terms of its usefulness and ease of use.

Anecdotal experience suggests that some requirements' specifications could be shrunk by between 25 and 50 per cent by the removal of 'wish list' items, bringing cost savings and a lower cost of ownership, as well as easier user adoption. It's important for the project team and sponsors to be able to define what business problem the ELN will solve and to ensure that user requirements are kept simple and focused on solving the problem.

Business case development and project management

Building a good business case requires a thorough and systematic approach to understanding current limitations and future requirements for the business. It is important to see laboratory informatics as a component in a lab ecosystem (technology, processes and people), rather than 'just another laboratory application'. The following points should all be considered in formulating the case for a new informatics system:

- Why do we need a new system?
- What is the problem that needs to be solved?
- Is there any quantitative data that illustrates the problem?
- Which laboratory areas will be involved in the project?
- Who makes the go/no-go decision?
- What are the issues relating to IP (internal/legal/patent)?
- Are there any regulatory compliance requirements?

Clarify why the organisation thinks it needs a new system. This is best achieved by developing a problem statement that quantifies a specific problem, or set of problems, about the laboratory's productivity and or knowledge management performance. The scope and scale of the problem (and hence, the solution) should be identified. The key decision makers/budget holders should also be identified, plus any other interested party who may have influence over a go/no-go decision. It is important to know what business level constraints may apply in terms of internal, legal or regulatory compliance.

Laboratory/company background:

- Use organisation charts to clarify roles and responsibilities and organisational relationships;
- Identify the nature and scientific disciplines of the laboratory work and how they relate to each other;
- Are outsourced agencies (contract labs) involved?

Establish the way in which the laboratory is organised, the nature of the work it undertakes and how it relates to internal and external organisations with whom it collaborates.

Current laboratory processes and systems:

- Which laboratory systems are already in use? (Are there SOPs?)
- Which data acquisition systems are already in use?
- Which teamwork/collaboration systems are already in use?
- Which document management systems are already in use?
- Who is responsible for the management and support of these systems?
- Is there a (electronic) records management policy?
- Are there any specific policies and restraints relating to the introduction of IT systems?

Establish how the laboratory is currently working, paying specific attention to the use and effectiveness of manual systems such as worksheets, paper lab notebooks and data management. Also identify major electronic systems used for the acquisition, processing and management of data, and ask what happens to this data, where is it stored and for how long? Is it communicated or transferred elsewhere – if so, how? Is it backed up and/or archived? Can it be found?

> **"It is important to know what business level constraints may apply in terms of internal, legal or regulatory compliance"**

Furthermore, consider if laboratory data is the responsibility of the lab, or does IT have any involvement, and what level of involvement does IT have in the purchase and implementation of laboratory systems?

Future laboratory processes and systems:

- Based on interviews with laboratory managers and staff, formulate a model that illustrates the major relationships between lab data and information;
- Construct data workflow and laboratory process diagrams;
- Identify any conflicts in nomenclature and establish an agreed taxonomy;
- Identify the role (scope and scale) of existing laboratory systems in the model and diagrams;
- Test the model and diagrams against each of the lab areas and other interested parties (IT, legal, QA, records management).

Put together a high-level plan showing the relationships, processes and data flows that describe a future state for the laboratory. This should include an identified role for each of the lab systems and should clarify the specific functions of each. Any problems with terminology should be resolved and the plan should be tested by presentation and discussion with the interested parties.

Business plan development:

- Quantify the benefits of the proposal, in particular productivity gains, ROI and knowledge management, and support these estimates with case studies;
- Undertake a risk assessment, paying attention to process, technology and people-related risks. Align the risk assessment to the set of user requirements;
- Prepare, and include in the business case, a high-level implementation plan that addresses any specific requirements and/or risks that have been identified.

Quantitative benefits should be identified, along with all risks. An implementation plan should address known risks and/or potential problems, in particular the strategic approach to roll out, e.g. a progressive deployment, the composition of the project team, change management and user support.

Human factors:

- What practical problems do laboratory workers experience with existing lab processes and data workflows?
- How well will laboratory workers accommodate change?
- Are there any cultural, political or other internal relationships that could have an impact on the project?

Identify potential problems associated with change management. This may be at an individual level (early adopters vs. laggards) or at an organisational level: R&D vs. legal, etc.

Internal culture and technology adoption

The introduction of multi-user IT systems into organisations has a mixed track record. Multi-user systems are usually specified by a project team and often contain a number of compromises and assumptions about the way people work. High level business objectives can therefore be put in jeopardy if users do not successfully adopt the new system. However, most case studies on electronic laboratory notebook (ELN) implementations indicate a positive user take-up. This may be attributed to the growing understanding of aspects of technology adoption, originally reported by Everett Rogers in his book *The Diffusion of Innovations*[17] and developed further by Geoffrey Moore in *Crossing the Chasm*.[18]

Moore's chasm (see Figure 7) is the gap between the early adopters and the mainstream market. The early adopters are a relatively easy market. Targeting them initially is important, but the next phase of the marketing strategy must target the conservative and pragmatic majority. The early adopters can play a central role in this. Since the ELN project team is likely to be formed from the early adopters, it can play a pivotal role not only in specifying and selecting a solution, but in articulating the rationale for the ELN, provide training and ongoing support to the conservative and pragmatic majority.

User adoption is often considered one of the most critical success factors of an IT project and paying appropriate attention to user requirements will enhance the likelihood of success. Key to this is the recognition that people are more likely to comply with a request when:

- A reason is provided;
- There is give and take;
- They see others complying;
- The request comes from someone they respect or like;
- The request comes from a legitimate source of authority.

Concerns about user adoption can be reduced by carefully choosing the project team to ensure that these criteria are addressed, rather than just announcing a new system and the training course schedule. Typically, putting a strong emphasis on user requirements and user adoption by engaging users throughout the process tends to brand the implementation as a 'laboratory' project, rather than an 'IT' project and this can often make it easier for scientists to accept the proposed change.

The Technology Acceptance Model[19] (see Figure 8) is an information systems theory that models how users come to accept and use a technology. It suggests that when users are presented with a software package, a number of factors influence their decisions about how and when they will use it. The main ones are:

Perceived usefulness (PU): 'The degree to which a person believes that using a particular system would enhance his or her job performance'.

Perceived ease-of-use (EOU): 'The degree to which a person believes that using a particular system would be free from effort'.

The technology acceptance model assumes that when someone forms an intention to act, they will be free to act without limitation. In the real world there will be many constraints, such as limited ability, time constraints, environmental or organisational limits, or unconscious habits which will limit the freedom to act.

Concentration on the positive aspects of 'usefulness', both to the organisation and to the individual, and 'ease of use' will help users develop a positive attitude. It is in this area that the early adopters can have a powerful influence of their conservative and pragmatic peers.

Fig. 8: Technology Acceptance Model

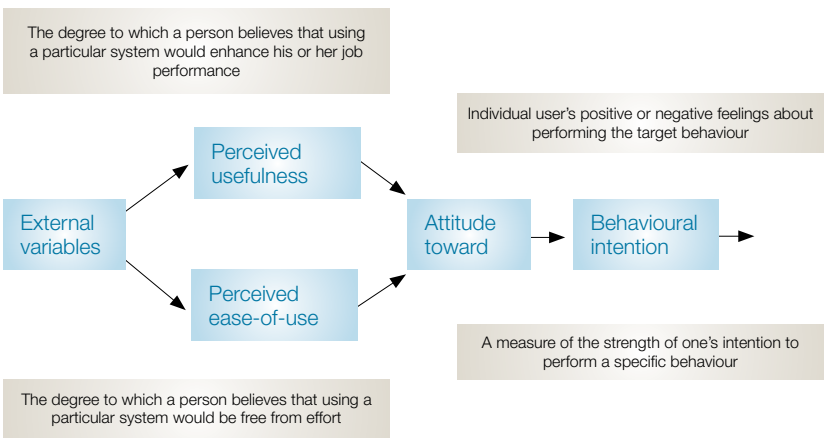
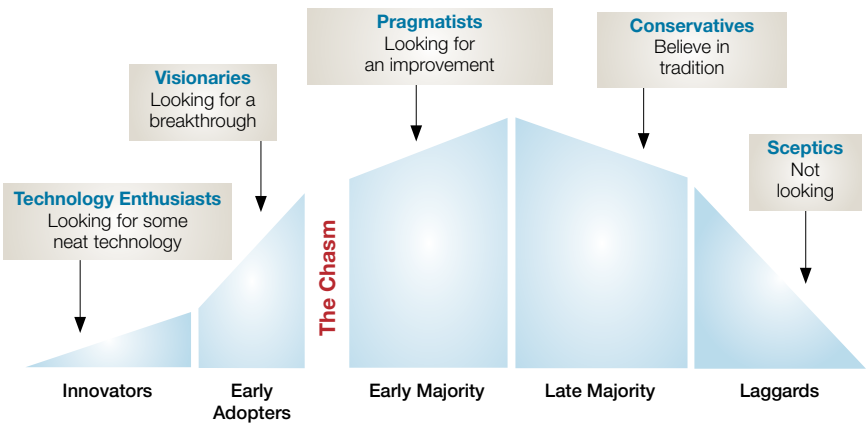


Fig. 7: Crossing the Chasm



Summary

The concept of a smart laboratory will vary from organisation to organisation depending on the nature of its business and the technological choices it makes. Discovery and development are increasingly recognised as two steps in a holistic product life-cycle process rather than standalone functions.

Innovation itself has moved on from 'Eureka moments' and chance discoveries to become a managed industrial process with an in-built need to address quality, regulatory, health and safety, and IP requirements. Just doing the science isn't enough anymore.

With this increasing demand for competency in science, technology and process understanding, an area of concern is how well adapted we as laboratory workers are to fulfilling industry's needs. The view is often taken that the demographic problems associated with 'technology literacy' will be resolved by attrition; new laboratory workers joining the organisation will have grown up in a digital world (digital natives). To some extent this is true, but technology continues to evolve at an ever-increasing rate and today's digital natives may find themselves challenged to keep up with further advances in technology, in the same way today's digital immigrants are.

It's not just a basic competency in technology skills that is needed, but also a deeper understanding of the continually evolving strategic and tactical roles that technologies play in the laboratory, both in terms of the science and the processes. Understanding how to use technology is only part of the answer; understanding why and ensuring that the right technologies are applied to the right processes is another matter.

Building a smart laboratory is dependent on three major criteria: deploying the right technologies, developing the right laboratory processes and creating the right culture. Keeping these three criteria in balance is fundamental to success. Perhaps the most critical issue is that of culture. Unfortunately this cannot be created or forced. As much as management may expect or demand their workforce develop a successful workplace culture, all that can be done is to create the right environment that will allow the culture to evolve. This can take a long time and, unfortunately, can be destroyed in minutes.

The focus of the guide has been on technology, in the form of laboratory informatics, with due consideration to the lab processes to which it can be applied. We have touched on some aspects of culture and technology adoption, but it must be remembered that almost every systems project defines user acceptance as a critical success factor. Technology on its own cannot do it.

References and further reading

References

1. Scott Jenkins, Pittsburgh Conference presentation on Paperless Laboratory, 2004
2. CENSA: The Collaborative Electronic Notebook Systems Association
3. Atrium Research: www.atriumresearch.com
4. The Gartner Hype Cycles: www.gartner.com
5. Good Automated Manufacturing Practice Guidelines version 5, International Society for Pharmaceutical Engineering, Tampa FL, 2008
6. McDowall, R.D., *Spectroscopy Focus on Quality*, July 2009, p23
7. *Using Electronic Records in Patent Proceedings*, article by Damien McCotter and Peter Wilcox. Originally published in Managing Intellectual Property's *World IP Contacts Handbook*, 14th edition, 2007. Available at www.mondaq.com
8. IP Expert Advice: Tips on creating a lab notebook that contains 'convincing evidence', www.edn.com/article/CA6445886.html?industryid=47048
9. Admissibility of Electronic in Interferences, Bruce H. Stoner Jr., Chief Administrative Patent Judge, www.uspto.gov/web/offices/com/sol/og/con/files/cons119.htm
10. Private communication: Colin Sandercock (Perkins Coie LLP) September 2011
11. The ABCs of Electronic Signatures, David Nettleton, *Lab Manager Magazine*, 9 September 2010: www.labmanager.com/articles_pf.asp?ID=707
12. The AnIML Programme: http://animl.sourceforge.net
13. The Pistoia Alliance: www.pistoiaalliance.org
14. SiLA Consortium for Standardisation in Lab Automation: www.sila.coop
15. International Consortium for Innovation and Quality in Pharmaceutical Development: http://iqconsortium.com/index.htm
16. Institute for Laboratory Automation: www.institutelabauto.org
17. Rogers, E. M., *Diffusion of Innovations*, The Free Press. New York
18. Crossing The Chasm, G.A. Moore, Capstone Publishing
19. Bagozzi, R. P., Davis, F. D., & Warshaw, P. R. (1992) *Development and test of a theory of technological learning and usage*. Human Relations, 45 (7), 660-686.

Further reading and websites

Stafford, J. E. H., (1995) *Advanced LIMS Technology: Case studies and business opportunities*, Springer,
Christensen, C. M., *The Innovator's Dilemma*, Harvard Business School Press
Segalstad, S. H., (2008) *International IT Regulations and Compliance: Quality Standards in the Pharmaceutical and Regulated Industries*, Wiley-Blackwell,
McDowall, R. D., (1987) *Laboratory Information Management Systems*, Sigma Press
Laboratory Notebook Guidelines: BookFactory, LLC, 2302 S. Edwin C. Moses Blvd, Dayton, OH 45408. Available at www.bookfactory.com
Mahaffey, R. R., (1990) *LIMS: Applied Information Technology for the Laboratory*, Van Nostrand Reinhold
Nakagawa, A. S., (1994) *LIMS: Implementation and Management,* Royal Society of Chemistry
Sellen, A. J., and Harper, R. H. R., (2003) *The Myth of the Paperless Office*, The MIT Press
Franklin, C., (2003) *Why Innovation Fails*, Spiro Press
Kanare, H. M., (1985) *Writing the Laboratory Notebook*, An American Chemical Society Publication

Industrial Lab Automation:
www.industriallabautomation.com
The Institute for Laboratory Automation:
www.institutelabauto.org
The Integrated Lab:
www.theintegratedlab.com
The latest information, news and articles on LIMS:
www.LIMSfinder.com
The independent, non-commercial LIMS user's group:
www.LIMSforum.com
The online encyclopaedia for Laboratory, Scientific and Health Informatics:
www.LIMSwiki.org
Free online training courses on LIMS:
www.LIMSuniversity.com
***Scientific Computing World*:**
www.scientific-computing.com

Technical Information Bulletin

Thermo Scientific CONNECTS™ for the Paperless Lab

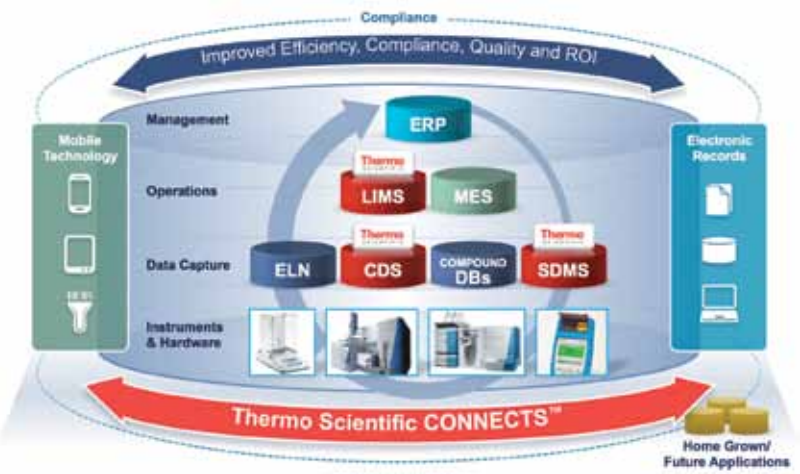
CONNECTS for the Paperless Lab is a combined *methodology, technology* and *services* offering that transforms laboratories into tightly integrated paperless environments.

- **Get real-time access to information**
- **Improve regulatory compliance and data integrity**
- **Automate processes and reduce manual data handling**

CONNECTS for the Paperless Lab delivers the *methodology, technology* and *services* necessary to transform today's laboratories into tightly integrated paperless environments, while leveraging existing investments in software systems, databases and laboratory instrumentation from a wide variety of commercial vendors.

By bridging the islands of data generated in the lab and transforming that data into information that can be used across the enterprise, organizations will reap the benefits of improved access to real-time information, regulatory compliance, data integrity, as well as time and cost savings by automating processes and reducing manual data handling.

CONNECTS for the Paperless Lab is comprised of three unique components — *methodology, technology* and *services* — that combined deliver seamless and continuous data flows from the lab into the enterprise, improving results



analysis, enabling faster decision making and improving productivity.

Building on a history of informatics innovations, CONNECTS provides access to all instrument data via a single interface and allows real-time analysis and investigation of results. The technology converts raw data to a vendor independent XML storage format to ensure future proof data archiving and facilitate data and information sharing across the organization, without having to rely on access to the original application. CONNECTS technology acts as a 'translator' of all the individual languages of the lab's disparate instruments and is able to deliver and accept data in the format appropriate for each instrument.

Thermo Scientific CONNECTS™ for the Paperless Lab:
Methodology: a comprehensive solutions-oriented consultation delivers an assessment of an organization's current level of laboratory automation, instrument integration and workflow. The assessment highlights potential ROI, operational gains and a plan for implementing solutions that address instances where current workflow relies on human intervention, which can be the cause of bottlenecks or introduce errors.

Technology: the vendor neutral suite of data integration, visualization and management technologies delivers comprehensive laboratory workflow automation, integrating instruments and equipment and connecting data sources with data destinations. The technology enables consolidation of analytical results, including access to raw data and instruments from web browsers and mobile devices.

Services: pre- and post-implementation consulting services are optimized to deliver current and future paperless environments, enabling real-time decision making based on access to aggregated information from the entire organization. CONNECTS Services are based on proven methodology for delivering successful projects using industry best practices.

20% REDUCTION

Did you know? By automating the laboratory companies save a tremendous amount of time and resources on manual activities in the lab. By automating the lab, CONNECTS for the Paperless Lab delivers improved efficiency, productivity, consistency, quality and reduced costs typically by over 20%.

For more information about CONNECTS for the Paperless Lab,
email us at marketing.informatics@thermofisher.com
or visit www.thermoscientific.com/paperlesslab

www.labware.com

SELECTING AN ELN?

Be sure to choose the right partner

Implementing a seamlessly integrated end to end laboratory informatics solution can truly transform your business, and leading such a project is a big responsibility.

LabWare's experience at developing and delivering enterprise-wide LIMS and ELN solutions is unsurpassed making us a great partner, whatever the scale of your project.

For over 20 years LabWare has been leading the way in laboratory informatics through continuous innovation and powerful results. Contact us to discuss how we can help make your next project a success.

Offices worldwide
supporting customers
in more than 90 countries

LABWARE
Results Count