Clarity. Vision. Enlightenment.™

Atrium Research & Consulting - 5 River Road Number 216 Wilton Connecticut USA 06897 - + 1 203 938 6924 - info@atriumresearch.com

# Smart Lab: Laboratory Informatics Exchange

15-16 February 2006

Knightsbridge, London, United Kingdom

## *Post Event Report*

**Prepared by:**

www.atriumresearch.com

Atrium Research is a scientific market research and consulting practice that provides the insight and education organizations require to cultivate successful informatics strategies and solutions. The company's mission is to provide best-in-class educational products on scientific informatics through balanced and unbiased information of the highest quality.

Atrium Research provides packaged and custom market research products and strategy, process re-engineering, vendor analysis, and technology readiness services. Atrium Research assists its clients to improve laboratory efficiency and financial performance through a structured method of process optimization and technology adoption.

# Table of Contents

# 1 Executive Summary

This report summarizes key themes discussed at the 2006 IQPC Smart Lab: Laboratory Informatics Exchange, held 15-16 February 2006 at the Carlton Tower on Cadogan Place, London, UK. The event included over 150 conferees from biopharmaceutical, chemical, government and consumer product organizations. In addition, over a dozen informatics software suppliers were present to discuss their solution suites. Leadership positions of attendees included vice presidents of information technology, directors of informatics, laboratory executives and managers, scientists, project managers, and system architects.

Speakers presented their insights from the selection and deployment of software solutions such as electronic laboratory notebook (ELN), laboratory information management systems (LIMS), integration technology, and scientific data management products. Workshops allowed for an open forum to discuss best practices and share lessons learned.

Several common themes were observed and noted during the course of the meeting:

- User involvement is absolutely critical to project success
- The majority of issues in deployment are user adoption, project management, and legal/regulatory rather than technical
- Merger and acquisition activity in the biopharmaceutical industry is complicating informatics strategies
- Point solutions developed during the last decade have caused a significant integration challenge
- The focus on Electronic Laboratory Notebooks (ELN) is increasing as a foundation architecture for a knowledge management strategy
- The strategy of "start small and build upon success" with an eye to an overall informatics vision is being used by more organizations
- Organizations are rethinking scientific IT governance practices and project prioritization strategies

Organizations who have successfully deployed informatics solutions are experiencing both tangible benefits (i.e. measurable improvements in laboratory efficiency) and intangible benefits (i.e. more insight into compound characteristics through integration of multiple data silos).

This document includes conference observations and dialogue themes on project drivers, considerations for successful deployments, an overview of the major technology discussed, and lessons learned.

# 2 Introduction

## 2.1 Background

The Smart Lab: Laboratory Informatics Exchange took place 15-16 February 2006 at the Carlton Tower at Cadogan Place in Knightsbridge, London, U.K. This meeting was an opportunity for leaders in the field of laboratory informatics to meet and discuss their experiences with the deployment of scientific software solutions. Topics chosen to be discussed included governance policies, user adoption challenges, integration strategies, new technology, intellectual property protection, knowledge management, and experience with specific technologies such data warehousing, electronic laboratory notebooks, and LIMS.

Speakers represented several industry types – biopharmaceutical, consumer products, chemical, academic, and government institutions - although the vast majority of speakers and attendees were from life sciences-based organizations, primarily involved in research and development. This is due to the significant challenges the life sciences face with managing vast collections of data and its exponential growth. However, the technological and user experiences of these organizations can be applied across a range of industry types.

## 2.2 Objective

The aim of this report is to capture and summarize key observations and themes expressed at the exchange. Confidentiality has been respected, so no individual or company designations have been included. The lack of attribution is not meant to slight the work or effort of any individual or company.

## 2.3 Laboratory Informatics in Context of the Exchange

Laboratory Informatics is a very broad concept dealing with the application of data and information technology to scientific processes and electronic record management. What technologies are used and how they are deployed is highly dependent on the environment (i.e. discovery, development, quality), the individual scientific domain (i.e. biology, chemistry, material science), and industry type (i.e. biopharmaceutical, chemical, food). For example, the application of a laboratory information management system (LIMS) in drug discovery is very different from how a LIMS is used for quality assurance / control. Where the LIMS in discovery must have significant flexibility to continually adapt to new experimental workflows, the LIMS in quality must be more rigid, enforcing compliance with a fixed set of procedures and methods. The advantages of a technology in one environment are often viewed as detrimental in another.

This is a significant challenge to informatics suppliers. To meet the demands of the user communities they serve, they are under constant pressure to adapt their technology to discrete application requirements - but at the same time attempt to do this across a wide range of industries. In many cases, this creates “deep” applications (versus horizontal or “broad” applications) that meet the needs of one department’s scientists, resulting in multiple point solutions needing to be deployed across the enterprise.

In the last decade, a typical organization has installed a plethora of these point solutions, creating a complex integration challenge as cross domain data needs to be assembled on compounds and projects. This issue of integration was a frequent theme throughout the conference.

The number of laboratory data types solutions must manage is noteworthy. There are several hundred data formats (i.e. instrument raw data files, spreadsheets, images, sequences, etc.) that exist and it is not atypical for a large enterprise to be challenged with the management, preservation and integration of dozens of varying record formats. In several sessions at the exchange, the lack of leadership on industry standards and exchange formats by suppliers was observed as a major source of user frustration. Though no solution was presented, the consensus was that pressure must be applied to suppliers to accept the few standards in progress, such as ASTM's AnIML and PDF/A.

Even with the expenditure of significant monies on informatics technology, organizations still face a remarkable use of paper. Paper records are used for the recording of experimental results in bound notebooks and for long term IP preservation. The reasons for this are many, but include: history / legacy, compliance, lack of data standards, and conservative legal departments. Companies looking to build an integrated informatics architecture and knowledge management strategy are realizing that paper-based processes are a considerable impediment to their vision. Electronic Laboratory Notebooks (ELN) technology and electronic record preservation were major topics of discussion woven into discussions on the use of e-records for intellectual property protection and regulatory compliance.

User adoption has always been a key to the success or failure of laboratory informatics solutions. Organizations are realizing – usually after failures of top down driven projects – that user involvement and acceptance is more important than the technology. Many speakers discussed their approaches to bridging the traditional gap between the scientific community and IT. Changes in governance practices, project management, and the movement to "bottoms up" approaches are being used by several organizations to diminish the risk of project failures caused by a lack of user adoption.

New technology such as predictive software, semantic integration, and pathway simulation is important to lower the costs of drug discovery and development and to streamline work processes. However, at the exchange, these emerging technologies were discussed only briefly as the challenges of integration, codification and paper elimination seemed to dominate conversations.

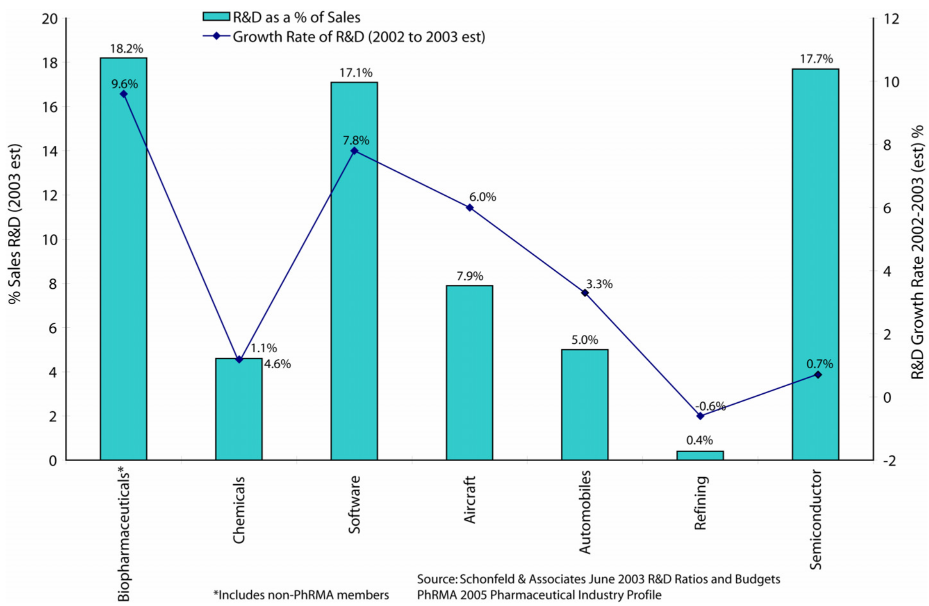
# 3 Drivers of Laboratory Informatics

There are many motivators of informatics projects, both external and internal. In this section, we will highlight macro drivers and the individual motivators expressed by those at the exchange.

## 3.1 Macro Drivers

- **Growth in data and complexity:** In the post-genomics world, data has not only been growing exponentially, but in complexity. Focus on new "omics" work is creating more composite datasets that must be assembled, analyzed and archived. Individual data silos must be seamlessly integrated to discover new data relationships and facilitate multiple query constructs.

- **R&D efficiency**: Organizations are under pressure to increase the efficiency of R&D. The biopharmaceutical industry in particular is forced to rethink discovery and development processes due to increasing expenditures which have not lead to a resultant growth in approved new molecular entities (NME). Informatics can assist in the facilitation of new processes and the breakdown of walls between individual departments.




Industrial R&D as a % of Sales and Growth Rates

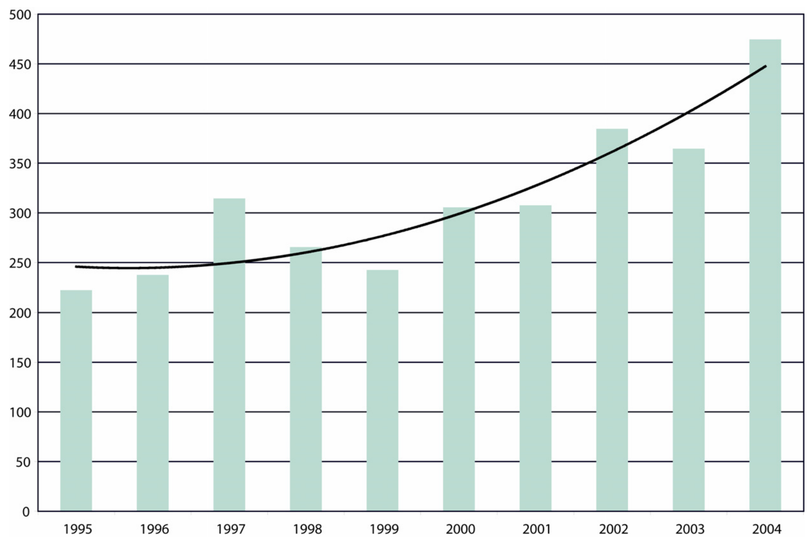
- **Personalized medicine**: The drive toward personalized medicine and the discovery of biomarkers is driving the need to assemble, mine, visualize and analyze data across multiple data repositories.

- **M&A activity**: The consolidation of life science companies is causing teams of researchers to work together like never before. Multiple geographic locations working on the same therapeutic area - or who have expertise in a particular field force - companies to re-think ways scientists communicate and share information. Project and compound data need to be integrated across multiple sites.

- **Safety assessment and data retention**: Possible changes in FDA guidance will result in storing and maintaining data on compounds long after they are in the patient supply chain, forcing the correlation of patient data with clinical and pre-clinical observations.

- **Knowledge Management**: Once data is reduced to a paper archive, it is rarely, if ever, reviewed again. The use of informatics to support a knowledge management strategy helps organizations achieve a useable institutional memory. Researchers can have a searchable repository of experimental information, find subject matter experts, and learn from the experiences of others. Failed experiments need not be repeated, improving efficiency. By have a consolidated knowledgebase, companies can look to re-purpose compounds and new uses of existing materials.

- **Competitive pressures**: Companies need to increase their agility and reduce the time to market for NMEs. This is not only against the generics, but also against ethical competitors to be the first to invent. Generics now comprise 57% of drug sales and this is expected to increase, so patent life must be extended.




FDA Generic Drug Approvals (Source: US FDA)

- **Regulatory:** Companies turn to informatics technology to improve their compliance with regulatory mandates, such as cGxP and Part 11. As the FDA develops new initiatives such as Process Analytical Technology ("PAT"), corporations are looking to automate and streamline operations while maintaining compliance.

- **Quality improvements**: The need to improve the quality of data forces organizations to look at software tools to institute compulsory standards and common data capture practices. Metadata must be managed in a consistent way to ensure searchability.

- **Faster decision making**: By having access to data and information, product decisions can be made in real time, versus waiting for collation, QA and review.

## 3.2 Individual Drivers

Conference participants expressed various drivers that were of particular importance to them:

- "Need to better manage resources"
- "Massive laboratory changes require data management flexibility"
- "Lab uptime must be increased"
- "Too many disparate point solutions"

- "Getting it wrong is not an option. We must manage information right the first time"
- "Users are frustrated with what we currently offer them. We must do thing differently"
- "Users are acting as data integrators"
- "We must capture intellectual capital"
- "Faster patent preparation"
- "We require rapid communication"
- "Improved archiving"
- "We need a single interface for multiple systems"
- "PAT is getting us to re-think our quality informatics architecture"
- "Automation of laborious manual tasks"

# 4 Considerations for Laboratory Informatics

Throughout the course of the conference, speakers presented various perspectives on project vision, developing a business case, organizational alignment, requirements definition, user adoption, and governance. This section will highlight many of those viewpoints.

## 4.1 Vision and Business Case

Projects must reference a clearly articulated vision for effective change management. Several speakers expressed their view that the development of an informatics vision and roadmap are critical to keep long term objectives in mind while tactically managing individual projects. However, a vision that was too broad and unrealistic was discouraged. Vision scope (as well as individual project scope) containment was seen as important to avoid runaway requirements and unobtainable project expectations. It was frequently commented that communication was an essential element of user adoption and project success. Therefore, the company's information vision must be clearly disseminated throughout the affected areas.

Vision questions to be addressed included:

- How does the informatics vision relate to the corporate strategy / vision?
- What is the improvement opportunity?
- How does the vision benefit the user community?
- What is the plan to complete the vision?
- What is the limitation of scope? How will it be maintained?
- What are the organizational requirements?
- Is the vision must be rational and achievable?
- How will a shared need be developed?

Speakers also offered their observations on the importance of a well developed business case. Consensus was the business case must tie back to the vision and provide additional detail including: objectives, benefits, return on investment, organizational impact, risk, costs and projected gains. The business case establishes sufficient details for a "go or no go" decision by management.

Speakers highlighted important considerations when developing a business plan:

- Definition of "what you don't know about what you don't know"
- Executive and front line management sponsorship
- Enthusiastic support from users
- Project sponsor must champion the business case through the approval process
- Alignment of benefits with organizational strategy
- Definition of integration points
- Regulatory requirements
- The baseline and performance metrics to measure success
- Definition of planning team and allocation of resources
- Benefits – i.e. time saved per week, capture of chemical reactions, etc.
- Return on opportunity, eNPV, payback, predicted outcomes
- Risks and dependencies
- A forum so users and stakeholders concerns can be addressed upfront, rather than later
- Change management, as project changes can occur during a project
- Quick wins and build from there
- Project team should be defined that is small and decisive. Avoid a large committee
- Should define who going to be accountable for the project success

## 4.2 Organizational Alignment

A number of presenters spoke of their thoughts on "how to bring IT and scientific organizations together at the management and laboratory level." It was noted that, in many cases, past informatics projects failed to deliver on promised – or expected – results due to disconnects between the information technology department and the scientific operations. In a show of hands, approximately 50% of the organizations represented at the exchange had placed scientific information systems under the control of the business operations to bridge the "culture gap." In this manner, the "bottoms up" approach of defining needs is believed to better serve changing scientific requirements. This "capability-based" organizational structure - rather than site-based – is said to offer added flexibility and faster decisions.

The globalization of informatics projects is having an organizational impact, as a several companies have assigned worldwide informatics leaders. Projects are being assessed, prioritized and resourced across multiple locations to more effectively address resource allocation. Gating and project pipeline processes are being put in place to judge projects based on their organizational impact and resource requirements (both human and capital). One company described their "Six Sigma-like" approach to project portfolio management.

In a few other instances, the chasm between discovery and development is being addressed through the appointment of a single head of R&D systems. It is believed that a single focus on cross domain informatics will help to transfer compounds from research to development in a more expeditious manner.

## 4.3 Requirements Definition and Planning

It was agreed among many of the participants that proper attention to requirements definition and project planning are critical to project success. However, many at the conference felt that often times requirements planning "goes too far," defining every last detail. Though the consensus was that "the devil is in the details," there was debate about which details are important. Process definition, data flows, use cases, and functional requirements were deemed important, but specifying the layout of very screen was not - at least for a COTS product.

Scope management was frequently mentioned. At least two presenters described systems that had unachievable requirements, causing them to waste both time and money. In at least one other case, the presenter described requirements that were so well defined, they were obsolete by the time they were completed as the needs of the business had changed.

Strategies for success were noted. These included: the use of rapid prototyping in partnership with the users, short term evaluations of supplier products during the requirements phase, and re-engineering and streamlining processes, and the creation of requirements to the "to be" state. "Start small" was most often cited as a key.

Other points made by speakers:

- Must have a clear understanding of the business information requirements
- Must have a clear understanding of the data structures
- Requirements will "snowball" without proper management and attention
- Requirements drift is inherent
- Should be viewed as a partnership between IT and the business
- Establish processes where none previously existed and link processes with business requirements
- Avoid being too vertical – keep in mind the vision and cross-functional requirements
- Project teams must include stakeholder representatives when developing requirements – but don't try to do everything
- Don't ignore the security requirements!
- Must have a functional perspective, rather than individual perspective
- Focus on immediate need, but keep in mind the long term
- Agree on definitions and ontologies upfront
- Requirements must be compatible with record management, QA and legal requirements
- Processes must be captured by the system – who is doing what, where they are doing it – what is the environment
- Can get overwhelmed by sophisticated requirements. Sometimes a simple approach can solve a majority of the needs
- Make sure to define the metadata requirements
- Define the change management process early
- Strive for consistent interfaces across projects
- Develop high level process maps, looking for opportunities. Agree on priorities
- Don't be afraid to throw projects away when necessary
- Make sure business rules are defined, especially with integration

- Define what types of data need to be managed and any data collaboration requirements
- Do any formats need to be converted?
- How will e-signatures be used? If not now, plan for the future
- Focus on the tools that provide benefit
- Make there is a plan for SOP writing, validation and training
- Plan for learning curve. There will be a short term decrease in productivity
- Plan for continuous improvement and change control
- Include a plan for upgrades and plan for on-going maintenance

## 4.4 User Adoption

In response to several surveys Atrium Research has conducted, the top answer to the question "Why do users resist using a software product" consistently is "General resistance to change." There is a bell curve of personality types when it comes to technology adoption (Moore's Technology Adoption Lifecycle) and the majority (80%) are either pragmatic or conservative. The pragmatists want to see what the technology can do for them and how can it help them in their daily life. The conservatives get on board only after all the pragmatists have adopted it. Change needs to be accepted by users and they must see how it can help them be more effective in their position. In case of scientists, it is often one of freedom from mundane tasks permitting more time for scientific endeavors, or software tools that provide new insights.

General Electric, in their process for change acceleration, defined project effectiveness as "E = Q x A" where "E" = effectiveness, "Q" = quality of the solution, and "A" = acceptance. Many first-class technical approaches have failed due to low user acceptance. Projects that do not take into consideration the human factor result in a sustained lower level of productivity, user resentment, and long term mistrust. GE uses a seven stage process to transition from the creation of a shared need to continuous improvement. This process is not too dissimilar to the various approaches discussed at the exchange.

Organizations have learned that early user involvement is not only essential to effective requirements definition, but it also reduces the risk of non-adoption. Users need to feel involved, communicated with, and have a connection with the project. Many of the rapid prototyping and user software evaluation tactics described at the exchange help to facilitate a relationship between the user and the project early on.

Speakers at the exchange discussed several approaches:

- Always keep in mind that the project is about people, not technology. The technology is there to assist the users or to support new processes, not get in their way
- There needs to be active user support and cheerleading from the user community
- Market the project internally
- Communicate, communicate, communicate
- Start small, demonstrate success and promote the success
- Peer pressure is much better than mandates from on top
- Front line managers need to be seen actively supporting the project
- Make sure the user interface is user-oriented. Must have as few screens and clicks where possible
- Make it easy to share data

- Highlight the projects added value and benefits
- How technological savvy are users? Assess and plan on appropriate levels of training
- Approaches to training are important. Train the trainer is best
- Integration is often a key. Show that fewer systems need to be dealt with
- Do your performance testing on the system and the infrastructure. Even if it's an infrastructure issue, the users will blame the system and it will get dropped. Make sure performance levels are consistent despite increasing loads
- Make sure the system adapts to the needs of the domain
- Communicate expectations lower than you think. It's always better to exceed expectations than the opposite
- Plan on a learning curve
- Make sure it works the way the scientist works
- Make sure it handles the data types the user needs

## 4.5 Legal

There was a lack of consensus about the need for paper to support a patent. Several large pharmaceutical companies who are deploying ELNs expressed their commitment to a "paperless" environment and plan to support patents based solely on electronic records. Others – citing their conservative legal counsels – are implementing ELNs with a "hybrid" model. In this scenario, paper notebook pages are printed, signed and bound for long term retention. Many agreed that the lack of consensus on the electronic IP issue has complicated their informatics strategies.

The reason for the confusion is that a laboratory notebook is often used as the primary evidence to prove inventorship and the proprietary rights that come with a patent. Details of the research and the dates and times of discovery and reduction to practice are critical to establishing these rights in the United States. This is because in the US, versus other countries, patents are awarded on a "first inventor" basis rather than on a "first to file." This puts the responsibility in the hand of the researcher to have accurate and corroborated records.

Many organizations believe paper records are safe. However, lacking or un-enforced notebook policies, poor record management practices, lack of witnessing and author signatures, and slow reduction to practice has prevented organizations from winning patent interference cases despite their use of paper records. Proper practices must be in place, regardless if the record was on paper or in a database.

Either paper records or electronic records are admissible in US courts under the US Federal Rules of Evidence (FRE). According to FRE, "[I]f data are stored in a computer or similar device, any printout or other output readable by sight, shown to reflect the data accurately, is an "original"." FRE 1006 permits the submissions of summaries of digital evidence in the form of "a chart, summary, or calculation" subject to certain restrictions. This is, of course, if the authenticity of the records was satisfactorily established. Hybrid ELN implementations, therefore, are not needed – a printout of an ELN e-record would be acceptable for admission as evidence if its reliability and record practices can be corroborated.

One speaker presented an approach to stamping their electronic records for corroboration. They transmit records to IP.com which applies a unique digital signature and time stamp from a trusted source to each record. This technology, developed and supplied by Surety, LLC, creates a unique digital signature which is the equivalent of a witness stamp. This ensures the authenticity of the creation date and creates a fingerprint of the record. Any subsequent modification of the record will result in a non-compatible signature and will be detected.

Currently, about 30% of ELN installations have committed to an electronic-only environment, up from 10% just three years ago. As additional major companies eliminate paper, this trend will accelerate as smaller organizations will follow their lead.

# 5 Technology Discussions

Numerous technologies were discussed at the conference, but chiefly among them were integration solutions, electronic laboratory notebooks, laboratory information management systems, and record retention / content management. These will be discussed below.

Future directions in "intelligent tools" for chemometrics and bioinformatics were also presented. Neural networks, machine learning, automated method development, pattern recognition, and computer-aided organic synthesis were talked about in the context of helping scientists work more intelligently by mining large data stores and finding data relationships that were previously unimaginable.

## 5.1 Integration Technology

Integration was the "hot" topic at the conference, as companies are looking to consolidate data from fragmented transactional systems scattered across the enterprise. Frequently mentioned were strategies for bringing together chemical and biological data either in a single common database or in an integration network.

The evolution of systems to Service Oriented Architectures (SOA) is perceived as a positive trend. Integration via services and XML data formats dramatically reduces the long term cost of ownership. The "loose coupling" of applications offers a level of flexibility organizations need, as the process flow requirements of scientific operations changes over time.

Suppliers, however, were often mentioned as creating barriers to integration. There was a noted level of frustration that these barriers increase the costs and complexity of integration:

- Different, incompatible data formats.
- Lack common exchange formats (i.e. experiments)
- Different security architectures
- Different architectures (some web service, some thin client, others not…)
- Weak application program interfaces (API)
- Lack of open architectures
- Vendors not thinking "big picture" with too much emphasis on point solutions

The need for establishment of a corporate ontology - or at a minimum a set of defined terms – was viewed as essential. Mergers and Acquisition (M&A) activity was expressed as a challenge to an integration plan, due to organization changes and the need to incorporate new (or having to re-map) defined terms.

Future trends, such as semantically relating data (i.e. Semantic Web) were discussed, though this was not perceived as having an immediate impact on corporate informatics given the challenges of legacy systems and paper knowledge repositories. It is a notable trend that must be closely watched.

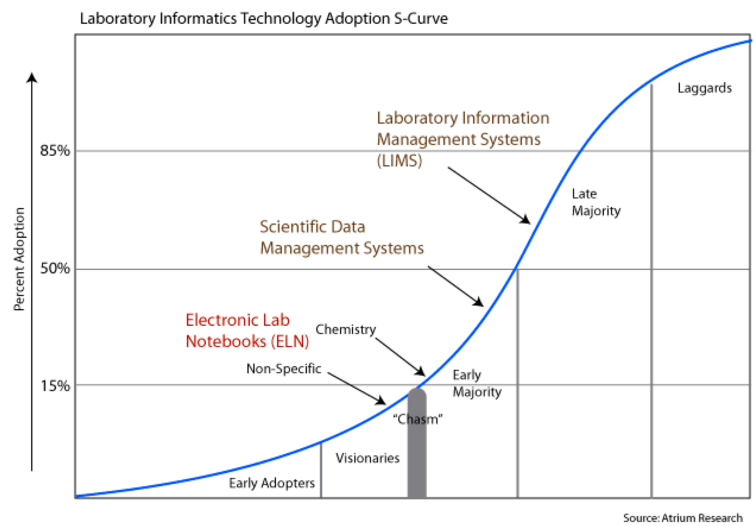
As with many aspects of laboratory informatics, there is no single solution for all problems. Discussed at the conference were integration approaches of:

- **ELN:** From the scientist perspective, the ELN can be a tool to integrate data from multiple systems (i.e. inventory, reaction database, bioassay database, compound registration, etc.) into a single interface. Though data may be posted into the individual systems, a single user interface simplifies the work of the user. In a sense, the ELN is acting as a portal, or scientists desktop, with the ability to also connect various data analysis tools.
- **Data warehousing:** Some organizations have, or are, exploring data warehouses to consolidate data from multiple transactional systems and use data mining and business intelligence tools. Downside is the ETL (Extract, Transform, Load) processing requirements, scalability and the need for a database administrator. The advantages are the speed of data access and technology maturity.
- **Federated:** In this model, searches are enacted across multiple systems with a consolidated result set. The complexity of implementing a federated model - given different data formats and security models - was discussed by several speakers. Federated searches can also be slow, but the architecture is highly flexible to add new data repositories when needed.
- **Hybrid:** One speaker, discussing the strengths and weaknesses of the federated and data warehouse models, chose to implement both solutions in areas where each best serve the needs of the business.
- **Ad Hoc:** New technology has appeared from companies in the last few years that are designed to join databases and files in an object-oriented workflow model. These scientifically aware tools are highly flexible for unique and specialized queries, data mining, and data analysis. However, their architectures are not viewed as being ready for enterprise-wide implementations.

## 5.2 Electronic Laboratory Notebooks (ELN)

Many speakers discussed their selection and implementation strategies for electronic laboratory notebooks (ELN). ELNs are in their early stage of market development and sales are growing over 30% annually. Only 20% of life science organizations have, or in the process of, deploying ELNs. The major area of growth for ELNs during the last three years has been in medicinal and process chemistry, but biology is beginning to take a strong interest.

There are many types of Electronic Laboratory Notebooks available and their use depends on the nature of the work performed. Replacing the drudgery of writing information into a notebook – cutting/pasting analytical results, photocopying pages, and microfilming – is just one benefit of an ELN. But an ELN is much more - an ELN is about the use of data and information and the creation of a knowledge repository. It is an instrument for improving resource effectiveness. Sophisticated ELNs include tools like chemical structure searching and drawing, reaction planning, experiment templates, integration with external research, dashboards, library integration, and embedded office applications such as Microsoft Excel.




ELNs are in an early growth phase

There are over 30 suppliers of ELNs. The various types of ELN products available are trying to solve different problems. For example, an ELN in research is designed to document the unstructured, free-flowing ideas and methods of the researcher while supporting a patent submission. In manufacturing, the ELN has more rigidity, as compliance with regulations such as GMP and 21 CFR Part 11 are important. This ELN variant is designed to capture compliance with established methods and procedures.

One presenter discussed their unique use of product lifecycle management (PLM) technology as a unique tool for integrating multiple scientific systems. This technology automates manual workflow tasks, manages records, and assembles documents and collects data from workstations and LIMS.

Speakers comments regarding ELNs included:

- Do I try and use one ELN with compromises or multiple that better address the needs of the scientist?
- The ELN has to be adaptable to the environment
- Integration is a key to success and adoption. It must not be "just one more system" the scientist has to use
- It must work the way the scientist works
- It must not slow down the scientist
- The interface must be clean and simple. If it is not easy to get data into it, scientists won't use it
- The ELN must provide very easy tools to get data out of it
- It takes 18-24 months for the ELN database to have sufficient information to be useful for searching
- Performance is critical
- There were varying opinions about the ability of an organization to adapt to an ELN
- Some organizations are viewing the ELN as an "Electronic Workspace"

- Workflow and information routing are important
- The security model must permit only those users who need to see data have access to it
- It may take 6-12 to change from paper to electronic documentation
- ELNs require enthusiastic leadership from the user community

## 5.3 Laboratory Information Management Systems (LIMS)

LIMS are technologies to manage laboratory workflow and structured experimental data. LIMS helps to organize and keep track of laboratory resources, samples, equipment and personnel. However, generic LIMS generally do not have the capabilities to manage unstructured records, experiment designs, analyze data, contextual data integration and annotation, or document assembly. Scientific Data Management (SDM), data analysis, spreadsheets, ELNs and other software complement LIMS to provide a laboratory informatics architecture. For some labs, a LIMS or an ELN is sufficient as a primary repository. This is dependent on the type of data to be managed and the laboratory processes.

Unfortunately, the majority of suppliers develop only one solution type (ELN, LIMS, SDM, etc.), creating an integration challenge for the customer. Several attendees commented that they would like to have suppliers approach the laboratory workflow problem with a fully integrated and complete offering.

## 5.4 Archiving and Record Retention

The "how and what" of electronic record archiving was discussed in the context of e-R&D initiatives. This is an commonly discussed topic at these conferences since there is so little consensus on an approach. It is up to the individual organizations to establish their own record retention practices and standards.

Electronic records must be preserved for a number of reasons, not the least of which is for the support of patents and for regulatory compliance. However, organizations are encouraged to take a risk-based approach and to not archive every record created in the enterprise. This "codify everything" strategy led to knowledge management initiative failures of the late 90s, due to thousands of unnecessary records being returned for simple queries (the "Google syndrome").

The lack of established archive formats was discussed. Portable Document Format (PDF) and the ISO standard PDF/A were the consensus format for document preservation. Microfilm and TIFF files are being used by some attendee organizations due to the uncertainty of future compatibility. However, the underlying issue of hundreds of incompatible analytical file formats – many of which support a patent and need to be maintained - was talked about in numerous sessions. Naturally, there is a strong desire not to maintain a "museum" of software, systems and operating systems just to reproduce a historical record. Conversion of some record formats to JCAMP and ANiML were discussed for long term preservation, though there can be no assurance of their long term compatibility. The ASCII nature of these formats, versus the binary raw data formats of supplier data systems, was seen as a major plus, even if formats change in the future. XML conversion appeared to be the preferred perspective of the majority, as conversion from one XML schema to another was not deemed to be of significant effort. Questions were raised, but not answered, about how to deal with – and validate - any data loss during file format conversion. It was talked openly of putting more pressure on the suppliers to address the matter of record formats.

Recording media was also discussed and the general agreement was that magnetic media was more stable as optical formats are in a constant state of change. Nevertheless, media migration strategies must be followed to compensate for media degradation regardless of media type.

# 6 Exhibitors and Sponsors

Participants at the exchange were able to discuss new products and approaches with the suppliers who exhibited at the event. The separate exhibition room and pre-arranged meetings provided an excellent format for one-on-one discussions and an educational opportunity for attendees. The following suppliers formally participated:

## 6.1 Co-Sponsors

- **IDBS**, a leader in biological data management, demonstrated several products for data management including their new BioBook Electronic Study Management (ESM) system and E-Workbook Electronic Laboratory Notebook (ELN)
- **Waters**, one of the largest analytical instrument companies, showed a suite of informatics products including their NuGenesis Scientific Data Management System (SDMS) and eLab Notebook ELN

## 6.2 Solution Providers

- **Adobe** highlighted their Intelligent Document Platform, which is a combination of Adobe Portable Document Format (PDF) and workflow/data exchange technology
- **Elsevier MDL** presented their various products for biological and chemical data management. Highlighted were the new MDL Notebook ELN, MDL Logistics, MDL Assay Explorer, and MDL Plate Manager products and the Isentris data integration framework
- **Informatica Software** showed their ability to address data integration challenge through their PowerExchange, Data Explorer, Data Quality and PowerCenter modules
- **Thermo Electron** illustrated their ability to manage laboratory information via their range of application-focused LIMS products such as Watson, Darwin and Galileo
- **UGS**, using their Teamcenter Product Lifecycle Management (PLM) software, showed how they facilitated an integrated ELN and IP Management environment at P&G and offered a vision for data/information integration

## 6.3 Exhibitors and Sponsors

- **CambridgeSoft** exhibited their desktop and enterprise solutions for chemistry and biology
- **KleeGroup** presented their consulting solutions and Kalabie ELN
- **LabVantage** highlighted their Sapphire Laboratory Information Management System
- **Tripos** demonstrated their Benchware suite including the Benchware Notebook ELN

# 7 Lessons Learned

One of the significant benefits of the conference was the sharing of knowledge. Many at the conference were keenly interested in users' experiences, both positive and negative. Below is a summary of presenters "lessons learned":

## 7.1 What to avoid

- "Committeeitis" Project teams getting out of control
- Trying to tackle projects too large of across all stakeholders
- Subjective prioritization
- Once a year review of projects
- Vision without a plan
- Too detailed requirements
- Disconnects between IT and the Lab
- Projects run in secret
- Top down decisions and directives
- Single person for requirements (guru)
- Culture as a reason for not doing something
- Avoid "Test sickness" – over stressing the users with testing
- Being too vertical in perspective
- Trying to store everything

## 7.2 What to strive for

- Leadership, not management. Leadership by example
- Small committees staffed by committed individuals. Push authority down
- "Just enough" requirements
- Be agile and make faster decisions. Have a sense of urgency
- Vision with prioritization and resource management
- IT populated with scientifically sensitive personnel
- Accountability
- "Bottoms up" – meeting the needs of the business
- Set expectations and meet them
- Agree on priorities
- Process. Identify workflows at different departments
- Strong user involvement
- Focus small, show success, move on
- Hands on testing by users
- Devil is in the details
- Open architectures
- Communication – keep message simple, but let users know what is in it for them
- Remove obstacles
- Short term wins