Implementation of an Enterprise E-Notebook: A Case Study

Simon Weston
IS Programme Manager at AstraZeneca

Introduction

The paper notebook has played a central role in the recording of the methods and results of scientific research for centuries. It has some strengths: portability, flexibility and (to some degree) incontrovertibility but in an enterprise environment it has many weaknesses. Chief among these is that the vast majority of information entered is lost to the enterprise unless substantial processes and governance structures are created to ensure it's dissemination. However, even if these are in place, searching for relevant information is likely to be time-consuming and difficult since there is no automated search capability below the level of the physical book.

In chemistry in particular, the relentless spread of automated, high-throughput synthetic approaches and the attendant growth in related analytical data outputs casts doubt on the suitability of the book format itself; there simply isn't room to accommodate all the data. Indeed, analytical data print-outs are commonly bound into a book form, signed and stored separately from the laboratory notebook itself.

The collection of paper notebooks holds the "corporate memory" and the intellectual property of the enterprise but both are likely to vary widely in quality from experiment to experiment. Cultural diversity and differences in standards, vocabulary and handwriting skill are almost impossible to reconcile and consistency is hard to achieve with ad-hoc, manual data entry.

This article describes the detailed justification for the project to implement an enterprise-wide Electronic Laboratory Notebook (ELN) solution for the Global Medicinal Chemistry community in AstraZeneca Discovery. It also describes the key components of such a solution and the approach we adopted for such a costly and complex project.

The Business Case for the Enterprise ELN

The benefits of an ELN system, exploited to the fullest extent, fall into two categories: tangible and quantifiable and intangible and qualitative.

The tangible and quantifiable benefits of the system are mainly related to personal productivity factors. For the AstraZeneca implementation, we estimate that an average improvement in the productivity of medicinal chemistry of 10% with a break-even point at 19 months post- deployment can be achieved. The scope of the implementation is likely to be a key factor influencing the magnitude and timing of the return on investment. In this case the scope was restricted to medicinal chemistry which, although practiced by almost 1000 people across Discovery, has a broadly consistent set of processes and work-flows associated with it.

The central argument for the tangible benefit is that an ELN increases the time available for conducting experiments by reducing the work-load of notebook entries and report writing. The following are some of the time saving features typically gained from an ELN:

- Cloning of replicate reactions - a chemist can make a copy of one of his own experiments to save time in writing up a similar one or can copy another chemist's experiment as a starting point

- Automatic calculations of weights, quantities and yields - many systems enable two-way automatic updating of computed variables

- Facilitating data handling for parallel chemistry - this is becoming increasingly important as high- throughput, automated chemistry becomes more prevalent

In addition, substantial time savings in preparing experimental sections for patents are possible, with all experimental records available in electronic form for easy incorporation into patent submissions.

Similar efficiencies should be observed in the preparation of reports and presentations, since a considerable portion of time spent on these tasks is spent finding and collating information, then transcribing it into a suitable format. In many cases, reference to the electronic notebook pages may be sufficient since the information is already present in a concise, legible format.

CambridgeSoft®
www.cambridgesoft.com

An In Depth Implementation of an Enterprise E-Notebook: A Case Study

A very substantial expected benefit is the creation of a searchable knowledge base of scientific experiments. This will tend to increase the value of experimental preparations and results (both failed and successful) by providing instant access to all chemists' data at any time.

are even anecdotal examples of experiments being repeated because that is quicker than searching for the relevant information. An ELN also promotes teamwork by providing access to the team's data.

The protection of Intellectual Property (IP) is enhanced mainly through an improvement in quality. Although there is little difference in the security of the IP, the fact that no

Table 1. Component Productivity Factors

Individual Factors	Hrs/Month/Saved	Number of Scientists	Hours/Year
Drawing reactions	0.167	1,000	2,000
Performing calculations	1.667	1,000	20,000
Risk assessment	0.667	1,000	8,000
Procedure writing	1	1,000	12,000
Registration of product(s)	1.667	1,000	20,000
Signing & Witnessing experiments	0.056	1,000	667
Indexing	1	1,000	12,000
Repeating experiments	1	1,000	12,000
Importing data	2	1,000	24,000
Monthly report/mtg prep	2	1,000	24,000
Experimental write-up for patents	3	1,000	36,000
Referring to past (archived) notebooks	0.7	1,000	8,400
Enterprise Factors	**Hours/Incident**	**Incidents/Year**	**Hours/Year**
Elimination of experiment failure	23.04	1,000	23,040
Collaborative problem solving	100	2	200
New employee ramp time	80	6	480
New project team member ramp time	10	6	60
Patent, regulatory attorney requests	5	10	50

The intangible benefits include the facilitation of cross-site project working which is something that is extremely inefficient or even non-existent because of the enormous cultural and practical barriers to sharing information in a large company. Information and data tend to reside on islands where proprietary formats and local practice preclude discovery by interested outsiders. There

handwriting is involved in the recording of the data and that business rules are enforced by default means that the quality is consistently high. Additionally, electronic audit trails and digital certification provide ample scope for demonstrating authenticity.

Implementing an enterprise ELN opens up a number of opportunities to integrate other information systems with

CambridgeSoft®
www.cambridgesoft.com

An In Depth Implementation of an Enterprise E-Notebook: A Case Study

Simon Weston
IS Programme Manager at AstraZeneca

experimental information. In the AstraZeneca implementation we have provided access to registration services from the ELN interface. We have also provided simple look-up type access to the corporate compound collection and some local reagent inventories to enable a more efficient experiment design workflow. Having an enterprise ELN holds out the promise of integrating with many types of scientific data though this should be addressed in a managed way by considering the business case and architectural design for each significant piece of work.

An ELN also enables the supervised capture and incorporation of hazard information so that synthetic procedures are performed under appropriate SHE conditions.

One final benefit is related to what is often referred to as "employer of choice": quite simply, scientists looking for employment want to work with the latest state-of-the-art equipment and tools, including software.

Table 2. Total Productivity Factors

Total hours/year	202,897
Total $/year	15,825,940
Productivity Correction Factor	0.6
Saved time put to productive use (hr/yr)	121,738
Productivity increases converted to cost savings ($/yr)	9,495,564

The Productivity Correction Factor accounts for the inefficient transfer of time saved to useful output and certain assumptions underlie these estimates (e.g. average hourly salary, no. experiments per month etc.).

The cost saving must be adjusted for take-up and user efficiency gain which are influenced by the implementation strategy.

Project Strategy

The development of the business case for an ELN and the process of selecting the most appropriate product were run as an integrated activity. The business case enabled the definition of a set of high-level requirements that formed the basis of an "invitation to tender" that was sent to a group of leading vendors of ELN systems. An in-house system also participated in the evaluation which involved chemists from all 10 sites and concluded that CambridgeSoft's E-Notebook Enterprise was the best fit to the requirements. The CambridgeSoft E-Notebook product can be installed and run "out-of-the- box". In practice however this is unlikely to meet the longer term needs and the system must be tailored through a combination of custom code and configurable form templates. CambridgeSoft's E-Notebook is a client-server application with a centrally-hosted database that is continually updated from a large number of distributed client PCs.

We employed three key strategies in the project which were designed to mitigate the risks related to scope creep, business change and the significant geographical spread of the project team and user community.

The first strategy was designed to try and minimize the inherent geographical risk by structuring the work as a program and regionalizing the implementation, i.e. the development of the product and associated services were centralized whilst the responsibility for the preparation of the local environments to receive the product was devolved to regional projects. The regions were defined as UK (incorporating two UK sites, one site in France and one in India), Sweden (incorporating three sites in Sweden) and North America (incorporating two US sites and one in Canada). The Product and Service Development projects were UK-based as was the CambridgeSoft project. Training and awareness was divided between a generic strategy, plan and material delivered as part of the product development project that could be customized by the regional projects to suit local conditions

CambridgeSoft®
www.cambridgesoft.com

An In Depth Implementation of an Enterprise E-Notebook: A Case Study

The second strategy was directed against one of the biggest challenges to delivering the business case for a large system with many stakeholders: the tendency for the scope to inflate during the development stage and repeatedly push the delivery of a usable system into the future. This is easily understood when you consider the sheer number of different perspectives being brought to the design. We sought to counter this tendency by adopting a time-box approach to the development of the AstraZeneca-configured system. In short, the delivery date of each release was fixed in advance by agreement with the vendor and Steering Group and the scope was allowed to vary to meet it. In practice, the scope only varies in one direction and the risk of reduced scope in functionality for each release was managed by the ruthless prioritization of features within each time-box. The support of stakeholders for this approach is critical!

The third strategy was set to counter the entrenched culture where experimental information is personal until final conclusions are derived and published. The sharing of preliminary information, especially experiments that go poorly, is often discouraged. We recognized that the benefits to be derived from the implementation of a global ELN would require a cultural migration amongst chemists that values the sharing of knowledge for informed decision-making above the personal ownership of information. This meant that successful delivery of both the system and the business case would require active management of the changes to working practice. For this purpose we identified the role of Business Change Manager to lead a group of key change agents in the individual chemistry groups to ensure that local culture and process could be aligned to the new way of working. It is critical that this "cultural migration" is managed by the stakeholders through the whole lifecycle of the product that supports the change and not just for the lifetime of the project: indeed a key deliverable of the project is an ongoing change-management process in Medicinal Chemistry.

Serious consideration had to be given to the implications of the decision by the Steering Group to rely solely on electronic signatures. Because we decided to go down this route, we elected to treat the Long Term Archive component of the system containing the signed PDF documents as a validated system (i.e. compliant with the provisions of FDA regulation 21CFR11). This, and the fact that we were implementing a COTS solution led us to run the project in a validated mode, using our proprietary Quality Management System.

At the time of writing, the product is being deployed to over 800 users and early indications are that the burn rate of the project is within 10% of the estimates in the business case. Whether the productivity gains will be realized remains to be seen but we are optimistic. The elapsed time from the approval of the business case by the stakeholder community and the delivery of the first release to the users is 12 months

Components of the Enterprise E-Notebook

There is much more to an enterprise ELN than the user interface (though as this is the visible face of the system it is very important). All enterprise systems need a server that hosts the data repository, client PCs and/or laboratory-based computing facilities to enable the efficient use of the system. In addition, there are a number of business and system processes that provide a supportive environment for the ELN.

The choice of software (or software architecture) for the ELN application itself is of course of critical importance with regard to the acceptance of a solution. In a distributed, multi-tier system, there is generally a trade-off between the richness of the client and the performance of the system. A richly functional client (typically one that enables complex calculations and/or has many sections for different types of information) may need to cache data locally and write it to the central database periodically. Digital security is a critical issue where an organization opts for a fully electronic signature solution. The system must apply the signature of the witness to an experiment in such a way that the resulting document cannot be altered by an unauthorized third- party.

An ELN can be used to capture all the required information about an experiment and requirements for what is captured and in what format can be enforced programmatically. The requirement for the repository to be easily searchable means that data quality must be consistently high.

CambridgeSoft®

www.cambridgesoft.com

An In Depth Implementation of an Enterprise E-Notebook: A Case Study

Simon Weston
IS Programme Manager at AstraZeneca

Standard Operating Procedures (SOPs) and business rules have been used in this case to ensure that compliance with the AstraZeneca naming standards and conventions is maintained. In the AstraZeneca ELN, the chemist is presented with a template and selects from a controlled vocabulary when entering textual descriptions. The user is also able to import other information already held electronically such as analytical data read-outs and other information that may support a claim to intellectual property. The template design and the required content are set by the business rules which in turn were set by the project reference group on behalf of the business. AstraZeneca Medicinal Chemistry has sponsored the development of a Global Standard Operating Procedure governing the use of the system by individual chemists. Local variations of the SOP are captured in appendices. The Global SOP stipulates the procedure for the recording and storage of data using the ELN system.

The application and database ELN software runs on standard computer configurations - it does not require any specialized equipment or management software. The only factor that caused any issue was the potential growth rate of the database. This rate is of course dependent on how well defined (and how well followed) the system SOP's are. We were faced with some tricky calculations based on volume per experiment per user per time. The final estimate of 0.3 TB/year led us to select an Oracle database solution on a dedicated server and incorporating database partitioning for the efficient management of the expected volumes.

If you are going to move all your experimental knowledge and intellectual property into an electronic system then the risk of losing the server on which the bulk of it resides must be managed. A combination of backup (with secure storage of media at a remote site), data replication and application clustering may be used to minimise the risk of data loss and disruption to business operations. The design, sophistication of this solution depends on the implications of the Business Continuity Plan (BCP). The BCP is owned by the business and describes how the business will continue to operate in the event that the system is unavailable. The plan states how long the business can operate using the workarounds defined in the plan and the length of this period determines the design (and cost!) of a disaster-recovery solution. This requirement was addressed by building a standby solution at a geographically separate location and using Oracle DataGuard to maintain a faithful copy of the data in both locations. This solution is "minimal data loss" and allows for the switching of the client applications to the standby database in the event that the main server is unavailable for a significant period.

From a user perspective, the system must be convenient to access and use. To this end, those laboratories that were not already equipped with remote access to the chemists' office-based PCs were upgraded to allow this. Many locations were equipped with remote-serving units that can locate a remote PC on the network and provide access to it using login credentials.

Conclusion

An enterprise ELN is a complex and costly solution to a common problem in large scientifically-driven companies: low visibility of the value of experimental information beyond the immediate team and cultural barriers to the sharing of such information. However, the quantitative benefits of an ELN can be clearly identified and measured and provided the requisite cultural change is managed carefully by the business, they should be realized. Careful consideration should be given to the strategies used for selecting, designing and implementing an ELN. Factors such as the number of users, the geography, the scope of features and the degree of change implied are critical inputs to the risk profile of the project.

The product marketplace for COTS ELN systems has matured in recent years to the point where leading vendors are providing systems of sufficient quality to justify purchasing rather than building the solution. It seems likely that over time, ELNs will become commodity products in the Pharmaceutical and Biotechnology software markets and a wave of consolidation will no doubt follow: all the more reason to focus on consistency of process, speedy deployment and effective exploitation to accelerate the return on the investment.

CambridgeSoft®
www.cambridgesoft.com