SELECTION OF A SCIENTIFIC DATA MANAGEMENT SYSTEM (SDMS) BASED ON USER REQUIREMENTS

Joseph Edward Heyward II

Submitted to the faculty of the School of Informatics
in partial fulfillment of the requirements
for the degree
Master of Science
in Chemical Informatics (Laboratory Informatics Specialization)
Indiana University
August 2008

Accepted by the Faculty of Indiana University, in partial fulfillment of the requirements for the degree of Master of Science in Chemical Informatics (Laboratory Informatics Specialization).

Master's Thesis Committee

_____

Mahesh Merchant, Ph.D., Chair

_____

Narayanan Perumal, Ph.D.

_____

Joseph A. Turpin

ACKNOWLEDGEMENTS

A sincere wish of gratitude goes out to all who have supported me during these years in the School of Informatics. I want to say a special thank you to my lovely wife, Krishna, for continuing to push and support me and for also putting her future education plans on hold over the last few years to support my endeavors. Thanks to the Eli Lilly and Company Educational Assistance Program for your financial support, and all my supervisors for working with me to coordinate my class schedules around work responsibilities when the need arose.

To my thesis committee, especially Dr. Mahesh Merchant (Committee Chair and Advisor), thank you for agreeing to be a part of this thesis research project, and for all of the honest feedback and resources provided.

ABSTRACT

Pharmaceutical research/development and quality control laboratories are faced with prodigious amounts of data from a multitude of heterogeneous data sources in a compliant manner, as mandated by requirements from regulatory agencies like the Food and Drug Administration (FDA). This has forced laboratories to use electronic data/information management systems to capture and maintain this data. Although the use of Laboratory Information Management Systems (LIMS) has gained widespread acceptance, pharmaceutical laboratories still struggle with the idea of system integration and lab automation using electronic laboratory notebooks (ELN) and more specifically scientific data management systems (SDMS), due to a normal resistance to change and a need to protect their existing IT investments. However, a properly assessed and validated SDMS offers the most significant benefits in terms of data quality, compliance, costs, and standardization across laboratories and interface capabilities to collect data from disparate sources and store them in a database for easy retrieval and management. In this research project, a detailed analysis of the functional and non-functional requirements for purchasing a SDMS, in addition to the analysis of the functionality of several of the mostly widely known SDMS is performed to determine which is most suitable for use in the pharmaceutical laboratories surveyed. Finally, validation requirements for a SDMS and more specifically computer system software in general is detailed and explained.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

Term/Acronym	Meaning
SDMS	Scientific Data Management System
ECM	Enterprise Content Management
CDS	Chromatography Data System
LIMS	Laboratory Information Management System
Native Instrument Data	File data, Raw data, and Files
Processed Data	Print data, Report data, and Results
API	Application Programming Interface
ID	Identification
LDAP	Lightweight Directory Access Protocol
Req.	Requirement
ER/ES	Electronic Records/Electronic Signatures
COA/COT	Certificate of Analysis/Certificate of Testing
RS	Responsible Scientist
QA	Quality Assurance
QC	Quality Control
PPR&D	Pharmaceutical Product Research and Development
COTS	Commercial Off-the-Shelf
OJT	On-the-job training
ROI	Return on Investment
R&D	Research and Development
FDA	Federal Drug Administration
CFR	Code of Federal Regulations
EU	European Union
LAN	Local Area Network
cGMP/GMP	Current Good Manufacturing Practices/Good Manufacturing Practices
GLP	Good Laboratory Practices
GxP	A generalization of quality guidelines, predominantly used in the pharmaceutical industry
GUI	Graphical User Interface

CDMS	Content Data Management System - is a mechanism for storing information that can be retrieved based on its content, not its storage location
ADMS	Automated Data Management System
UNIX	A computer operating system originally developed in the 1960s and 1970s by a group of AT&T employees at Bell Labs
BLOB	Binary Large Object (in computer systems)
SDK	Software Development Kit
CIFS	Common Internet File System
NFS	Network File System
XML	Extensive Markup Language
HTML	Hypertext Markup Language
PST	A filename extension used with certain Microsoft Windows products
PDF	Portable Document Format
CAS device	Content-addressable storage
UML	Unified Modeling Language
GAMP	Good Automated Manufacturing Practices
BCP	Business Continuity Plan
DRP	Disaster Recovery Plan
PDA	Parenteral Drug Association
APV	Arbeitsgemeinschaft Pharmazeutische Verfahrenstechnik
UAT	User Acceptance Testing

CHAPTER ONE: INTRODUCTION &BACKGROUND

Introduction to subject

The exponential growth of data produced in biopharmaceutical laboratories today has forced the need for moving from capturing data on paper, or storing it in spreadsheets, and small, non-robust databases to the need for having an automated and secure data management platform. The increase in data volume has been accompanied by an increase in data formats for storage, organization, and dissemination of heterogeneous data [1]. In the November edition of the 2003 Scientific Computing & Instrumentation LIMS Guide, M. Elliott pointed out that, traditionally laboratories have looked to Laboratory Information Management Systems (LIMS) to assist in managing the ever increasing information workload. However, with new regulations and reporting requirements, this has brought about the development of scientific data management systems (SDMS), to not only collect data but also manage them in a way that insures long-term preservation and knowledge retention [2].

Scientific Data Management Systems

A scientific data management system (SDMS) is used to collect, organize, index, store, archive, search, and share electronic records. It provides a secure, central repository, and rich content services to allow organizations to manage and re-use business critical information, comply with regulatory and corporate mandates, and enable collaboration for any type of electronic record [3]. P. Kegelmeyer, et al, noted that the overall goal of scientific data management is to hide the complexity of the underlying technologies, thereby freeing the scientists to focus on data comprehension [4]. Also

they pointed out that to be effective, scientific data management must provide an environment that allows scientists to store retrieve and search data within the natural context of their work [4]. Scientists need to be able to store data from different instruments and other document applications easily, find it quickly and share it with colleagues around the world.

Scientific Data Management Systems (SDMS) designed for the pharmaceutical industry, which includes such solutions as Waters NuGenesis SDMS (Figure 1) and Agilent Cerity Enterprise Content Management (ECM), are vital in this new era of scientific data management. They can solve many of the problems in data management on a large-scale and provide an expanded choice for scientists.

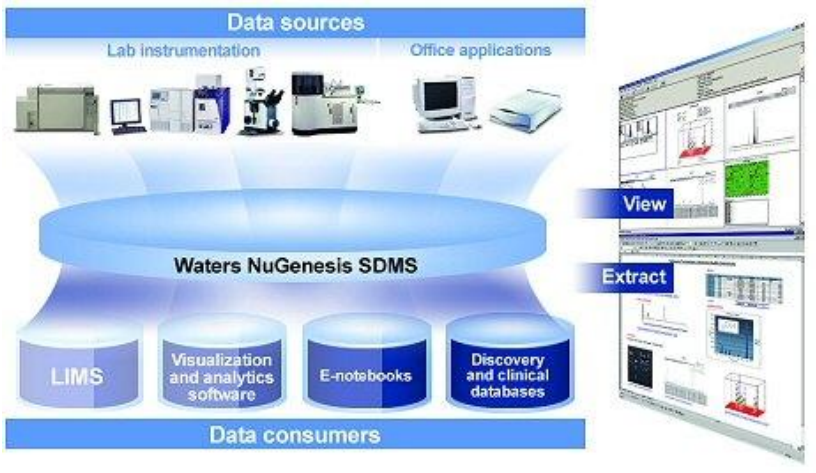


Figure 1- Example of SDMS

(Used with the permission of Waters Corporation.)

Laboratory Workflow

What is a laboratory workflow? Taken from the Wikipedia definition, "a workflow at its simplest is the movement of documents and/or tasks through a work

process. More specifically, workflow is the operational aspect of a work procedure: how tasks are structured, who performs them, what their relative order is, how they are synchronized, how information flows to support the tasks and how tasks are being tracked". The following is an outline of a general laboratory workflow:

1. Sample Preparation Phase
    - Sample collection and transport
    - Sample accession
    - Sample assay preparation
2. Sample Processing Phase
    - Creating a work list
    - Running assays
    - Test accuracy
3. Results Analysis Phase
    - Results validation
    - Repeat testing
    - Verification of false-positive and QC for false-negative results
    - Turnaround time efficiency
    - Reporting results (i.e. using LIMS, SDMS)

The pharmaceutical division of a global healthcare company can improve data and intellectual property management capabilities by implementing properly validated laboratory informatics solutions, particularly scientific data management systems (SDMS). SDMS validation is an important issue for many laboratories, particularly those

that operate in regulatory environments. The laboratory should develop a formal plan for validating the system, including the test data and acceptance criteria.

Computer System Validation

Computer system validation is a process that provides a high degree of assurance that a computer system is reliably doing what it is intended to do and will continue to do so. The goal is to produce a reliable system and that produces good data. The end result is documented evidence of the validated system in the form of a validation package.

Validation of computer systems and software used in pharmaceutical laboratories must comply with GMP and GLP regulations to ensure that they perform as defined by the functional/user requirements.

Validation is important for many reasons, but the three main reasons why a company should validate a computer system – more specifically a SDMS – are to protect their investment, provide consistent product quality, and to comply with regulations.

The investment in computer systems, including SDMS, has risen dramatically over the past years. Validation is a way of building quality into a computer system (e.g. SDMS) and increases the odds that the system will meet expectations [5].

Product quality, just like "customer" can be used in a wide scope. The product for a company is its patent, commercial product, whereas, the customer(s) are the consumers who purchase the product. The product of a laboratory is information and, as such, with research and development laboratories validation is used to ensure that the results generated to support product development are accurate [5]. Scientific data management systems, like chromatography data systems (CDS) and laboratory information management systems (LIMS), are becoming increasingly used in the

manufacturing environment; therefore, it is essential to know that the quality of the data that are entered into and extracted from a SDMS to support a product release is good, thus ensuring the overall quality of the final product.

Both the Food and Drug Administration (FDA) and the European Union (EU) expect manual and computerized systems to show equal quality [5]. Software validation is a requirement of the Quality System regulation, which was published in the Federal Register on October 7, 1996 and took effect on June 1, 1997 [6]. By validating a SDMS throughout the project life cycle using good validation practices, a company and/or laboratory could reduce the risk of non-compliance with regulations such as Part 11 of the FDA's Title 21 of the Code of Federal Regulations (CFR) that deal with Electronic Records and Electronic Signatures. Also following good validation practices helps to build a better rapport with regulatory agencies that perform periodic audits of these pharmaceutical companies, establishes a good foundation for better management control throughout a global company, and fosters better communication across teams.

An established comprehensive software validation process helps to reduce the long-term cost of software by reducing the cost of validation for each subsequent release of the software [6].

The full validation of a commercial-off-the shelf (COTS) system is very costly. For example, this would mean the testing of each function of the software under normal and unexpected conditions across the expected application range, and for each possible configuration of the system. Moreover, whenever the system changes, full revalidation would require identical tests be rerun. In today's rapidly changing computer environment, this could possibly mean that the system would be used 100% for testing

[7]. This is why "risk-based validation" is now a commonly heard phrase in the pharmaceutical industry. The risk-based validation approach will be discussed in subsequent sections.

Validation Life-Cycles

The FDA guidance on General Principles of Software Validation does not recommend a specific life cycle model, as these models should be established as appropriate for the product and organization. The software life cycle model that is selected should cover the software from it birth to its retirement [8].

A software product life cycle is defined as the stages the product goes through from its design to its decommissioning. Other stages included are construction, startup, production, and maintenance. A typical example is pictured below (Figure 2). In a poorly planned computer system implementation, this is not always the case. This could be due to several reasons, from a lack of understanding the importance of validation, to the additional costs that could possibly be added to the project for documentation, for laboratories working under regulations such as cGMP.

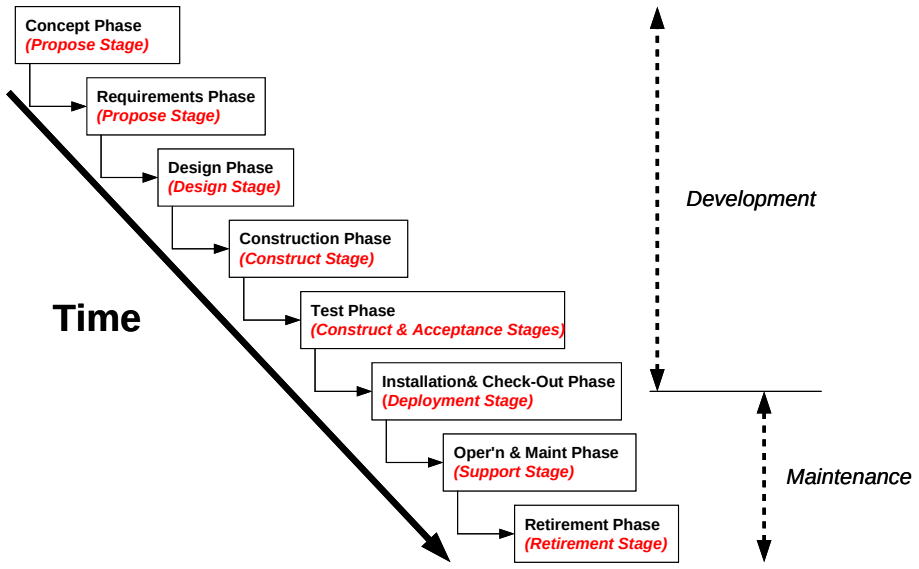


Figure 2: Software Life Cycle

Computer system validation issues have been addressed by several industry organizations and private authors over the years. The Good Automated Manufacturing Practices Forum (GAMP) developed guidelines (latest release December 2001) for computer validation. These guidelines are aimed at assisting companies in the healthcare industries to achieve validated and compliant automated systems. The Parenteral Drug Association (PDA) developed and published technical reports, with input from consultants, on validating computer related systems (Report 18-1994) and the validation and qualification of laboratory data acquisition systems (Report 31-1999) [9, 10]. The Arbeitsgemeinschaft Pharmazeutische Verfahrenstechnik (APV) also published guidelines, with input from regulatory agencies, based on its interpretation of Annex 11 EU Guide to GMP. An English translation is available in GAMP 4. L. Huber also published validation reference books for the validation of computerized analytical and networked systems (i.e. "*Qualification and validation of software and computer systems*

in laboratories"). The most notable of these for use in pharmaceutical companies is the GAMP lifecycle, which I have chosen to detail more in the next section.

GAMP

GAMP 4 categorizes software in one of the 5 categories below [38]:

GAMP Class	Category	Validation Action
1	Operating systems	Record Version
2	Instruments and controllers	Record configuration and calibration
3	Configurable packages	Audit supplier, validate any bespoke code. Apply full life-cycle requirements.
4	Systems where the entire code or part of the code are configurable.	Audit supplier and code, validate any bespoke configurations apply full life cycle requirements.
5	Systems utilizing custom which develop predicate rules information.	Audit supplier, validate all code, and apply **full life-cycle requirements.**

Due to the great variety of medical devices, processes, and manufacturing facilities, it is not possible to state in one document all of the specific validation elements that are applicable. However, a general application of several broad concepts can be used successfully as guidance for validation [11].

The following is an outline and explanation of each of the broad concepts/elements used for validating a computer system, based on regulatory guidance and industry methodologies such as GAMP:

- ✓ Validation Plan
- ✓ Vendor Evaluation and Management
- ✓ Requirements and Design (including ER/ES)
- ✓ Testing
- ✓ Change Management
- ✓ Security
    - o Physical Security
    - o Logical Security
- ✓ Business Continuity Plan (BCP)
- ✓ Disaster Recovery Plan (DRP)
- ✓ Periodic Review
- ✓ Training
- ✓ Validation Report
- ✓ Documentation Storage and Retention
- ✓ Decommissioning and Retirement

Validation Plan

- Must be written in the development phase.
- Based on the criticality and complexity of the computer system and business process.
- Defines the roles and responsibilities of the parties involved in the validation, including that of the vendor.
- Must be kept current using change control until the system acceptance is complete; then it becomes historical.
- Must be approved by the System Owner, System Custodian, and Computer System Quality before formal testing begins.

Vendor Evaluation and Management

- A vendor evaluation describes how the vendor (for COTS) was evaluated, the impact of the evaluation on validation activities.
- The goal is to determine if the vendor produces a quality product.
- The results and conclusion of the evaluation must be documented.

Requirements and Design (including ER/ES)

- Requirements must address intended use of the system, which may not cover all system features. See section of 'Requirements Gathering' for more detail on requirements.
- Must address functional, security, and ER/ES requirements.
- Must be kept current and be traceable to the design specifications and testing.

- Approved by the System Owner, System Custodian, Business Quality, and Computer System Quality (for GxP systems).

Testing

- Includes planning, execution, and summarization of testing activities.
- There are four levels of testing (Unit, Integration, System, Acceptance), with System-level being a minimum requirement for all systems.
    - Unit testing – a software development process in which the smallest testable parts of an application, called units, are individually and independently scrutinized for proper operation. Unit testing is often automated but it can also be done manually [12].
    - Integration testing – a software development process in which program units are combined and tested as groups in multiple ways. Integration testing can expose problems with the interfaces among program components before trouble occurs in real-world program execution [13].
    - System testing – is testing conducted on a complete, integrated system to evaluate the system's compliance with its specified requirements. System testing falls within the scope of black box testing, and as such, should require no knowledge of the inner design of the code or logic [14].
    - User Acceptance testing (UAT) – also called beta testing, application testing, and end user testing - is a phase of software development in which the software is tested in the "real world" by the intended audience [15]. UAT is one of the final stages of a project and often occurs before a client or customer accepts the new system [14].

- Details must be addressed in the Test Plan.
- Test scripts must be approved prior to execution.
- A process must be in place to handle system defects, code errors, and test script errors.
- The Test Summary report must address failures and any remaining open issues.
- The Test Plan and Summary report must be approved by the System Owner, System Custodian, and Computer System Quality (for GxP systems).

Change Management/Control

- Governed by a SOP that describes the method for controlling and communicating system changes.
- Changes must be agreed upon by the System Owner, all impacted areas, and Business Quality.
- Change control documentation must include;
    - The reason for the change
    - The impact of the change
    - Who made the change
    - Who approved the change
    - Results of the change
    - Date/time the change was introduced into production (in order to have a back-track point (if possible)).

Security

- Defines two main deliverables:
    - Security Plan
        - Includes physical (i.e. card reader) and logical (i.e. password) controls to prevent unauthorized access to the system.
        - Identifies risks and countermeasures.
        - It must document who has access and privileges to what and a revision history must be maintained.
    - Security Administration SOP
        - Defines the process for granting and revoking system access.
        - Includes the process for performing access roster reviews.
- Both the Security Plan and SOP require approval by the System Owner, System Custodian, and Computer System Quality (for GxP systems).

Business Continuity Plan (BCP)

- This is a business activity, not an IT activity, although it may involve assistance from the IT department.
- The intent is to protect critical business operations until the system is restored.
- Defines how the business will continue during a computer system outage.
- The process should be tested periodically and documented.
- A copy of the BCP, along with any necessary reference information should be maintained in a separate and secure location.

Disaster Recovery Plan (DRP)

- The intent is to assist the business/company in responding quickly and effectively to a disaster (i.e. fire, flood, unexpected hard disk failure or a power failure or even a hard disk failure or unexpected partition damage, or a virus attack).
- DRP should be based on an evaluation of risk of the type of computer system, organization complexity, impact of the quality and safety of the product, and who is providing support to name a few.
- The ongoing disaster recovery process secures, protects, and backs-up the data on some frequency (i.e. daily)
- The process should be tested periodically and documented. Testing is as much or even more important than simply creating a plan and leaving it at that.
- A Disaster Recovery SOP should be written to document roles and responsibilities during a disaster.
- A copy of the DRP, along with any necessary reference information should be maintained in a separate and secure location.

Periodic Review

- This is an assessment to assure continued compliance of the computer system.
- The review must include:
    - Changes made since the last review.
    - Deviation, system, error, and maintenance logs.
    - Software upgrades.
    - Open action items
    - Security roster and privilege review

- The report must be approved by the System Owner, System Custodian, and Computer System Quality (for GxP systems).

Training

- Documentation should be maintained to prove members of the project team, support staff, and end-users are properly trained and qualified.
- Qualification should be based on a combination of previous experience, education, and on-the-job training (including SOP training).
- As the computer system or the use of it changes, some level of re-training is required. The amount should depend on the complexity of the change.

Validation Report

- Describes the results of the validation plan activities.
- Explains any deviations from the validation plan.
- Includes a list of deliverables generated during validation and the location (electronic or paper form) of the deliverables. These include the validation plan, SOPs, user requirements, and system specifications.
- Concludes whether the computer system is fit for production use.
- Approved by the System Owner, System Custodian, and Computer System Quality (for GxP systems) before being put into production.

Documentation Storage and Retention

- Organized and retained based on regulatory GxP requirements per Corporate Record Retention schedules.
- Readily available for review, and to maintain validation, as it is a continuous process, as is noted in the section on "*Retrospective Validation*".
- Retained electronically or on paper with security access control to prevent unauthorized changes, destruction, or even worse, loss.
- A SOP must address how access is controlled, storage, and retrieval of validation documents.

Decommissioning and Retirement

- A change control or validation plan must:
    - Address data archival and/or migration (i.e. to a server or a new system).
    - Detail how long the data source code will be retained and where it will be stored.
    - Define a backup plan in case the need arises to return the system to production (temporarily or permanently).
- The change control or validation plan must be approved by the System Owner and Quality Assurance before officially taking the system out of production use.

Risk-Based Validation

Regulatory agencies (e.g. the FDA) expect a documented risk assessment for each computer system otherwise a full validation is required. In the past pharmaceutical

companies (and others) have often used this approach; however, there was minimal documentation to justify such an approach. Consequently, the approach was not implemented consistently within the company. The common impression it that this [risk-based] approach is a method that will reduce the overall time and effort expended in validation, and therefore will positively impact productivity and profitability [16]. Though this may be true for well-planned and executed risk-based validations, if there is a lack of documentation and thus a lack of knowledge of how to implement such an approach, chances are the real benefits will not be seen.

Risk-based validation involves two steps. The first is to define the risk category as high, medium, or low. The final step is to define the extent of validation for each category according to the guidelines laid out by the company, since there is no generally accepted model to follow, and no universal solution. Each company must figure out what level of validation suits them best, because success really depends on the company's unique situation.

Risk Assessment and Management

Risks that pharmaceutical companies deal with include patient risk (safety and efficacy of drugs), regulatory risks [FDA 483's, Warning Letters, product recalls, etc], and financial risk [7]. Below is a general overview of the phases involved in the risk management approach for validating COTS, and an example risk management evaluation documentation table.

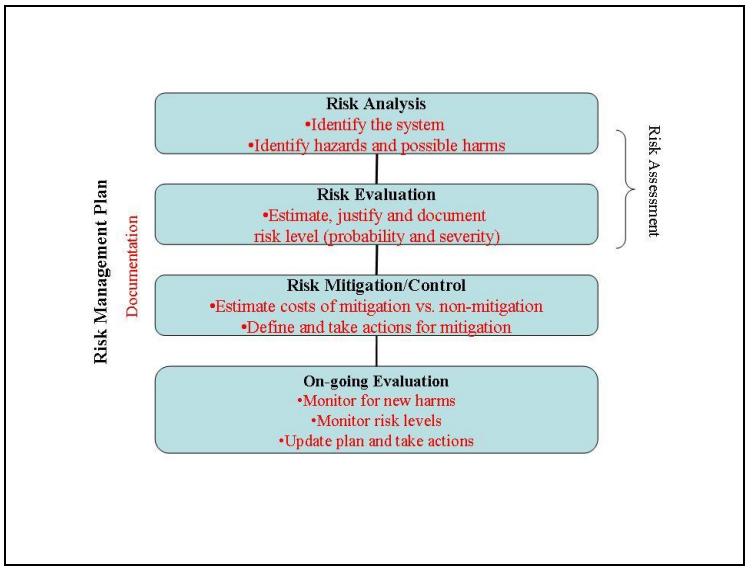


Figure 3: Risk Management Phases

Name/Organization: Date:		System ID:	Location:
Risk Description	Impact of Possible Harm	Probability of Occurrence	Method of Mitigation

Figure 4: Example of a Generic Risk Management Evaluation Documentation Table

Retrospective Validation

Retrospective validation is based primarily on reliable operation and proof of performance in the past rather than on qualification during development and installation. The exact amount of validation would depend on the complexity of the system and its current use. Therefore, the effort involved will vary from system to system [17].

The process discussed here is a general attempt to define a strategy for performing retrospective evaluation and validation of an existing system. Like performing a risk-based validation for a new system, there is no single method to perform retrospective validations.

The first step in the process is to define and document the current system use and user requirement specifications [17]. The next step is to collect information and documentation on the history of the specific system under evaluation, including the type and frequency of initial and ongoing testing but also the dates and installation procedure of updates to the system. Finally, the documentation should be evaluated in relation to the user requirement specifications (including functional specifications), and performance limits previously defined.

As mentioned in the previous section on validation, retrospective validation is based primarily on reliable operation and proof of performance in the past. It begins with the collection of system information such as test results, failure logs and maintenance records, and changes to the system over its use. It is up to the laboratory performing the validation to determine if enough tests have been performed to verify proper system performance as described in the user requirements. If not, additional tests should be developed and executed. If the system passes all the requirements, it should be treated as a new system [17].

Importance of subject

Data management is of critical importance in laboratory informatics. If the central task of laboratory informatics is to maximize laboratory operations (particularly in analytical, production, and research and development (R&D) laboratories) through the application of information technology, it is crucial that all associated data be accurately captured, annotated, and maintained, even in the face of rapid growth and frequent updates. It is also critical to be able to retrieve data of interest from multiple distributed heterogeneous data sources in a timely manner, and precisely enough to be able effectively to separate them from the distracting noise of irrelevant, unreliable or insignificant data. These issues are of concern to scientists themselves, and are of greater concern to non-expert users. Tools such as standards, visualization and analysis, will be critical to taking advantage of scientific data [18]. Scientific Data Management Systems can be a critical part of this equation.

Data quality is significant for a decision based on bad data is a bad decision. Data quality issues first arise during the initial application design stages when requirements for extracting and transforming data from operational systems are developed, and remain an ongoing concern throughout application development, use and maintenance. Poor data quality can impact an organization in many ways; for example the integrity of databases and other information systems become suspect. Achieving high data quality requires establishing some sort of departmental or organizational standards to help maintain data consistency and quality. This involves getting the people responsible for creating or manipulating the data, such as end users, entry people, etc., to assume a sense of ownership for data quality.

Data accessibility is equally important because if data are not accessible, it is the same as if they did not exist [18]. Today it seems unimaginable that data are not accessible, and large amounts of existing data are being ignored because of the costs associated with computerizing older, outdated paper data collection systems. Pharmaceutical companies must work together to make sure that investments in science are not constrained by lack of accessibility. Data themselves are rarely profitable; their use is. Consequently, the integrity of data is very important.

Due to the well-being of the safety of the product and consumer, data integrity is perhaps of most significance in the management of scientific laboratory data in a pharmaceutical laboratory. A SDMS, like other laboratory informatics solutions, must be capable of providing full versioning and audit-trails for all human readable documents and reports, including the MS Office applications Word and Excel. This feature provides the needed validation to ensure that the captured data is maintained in a compliant manner.

CHAPTER TWO: LITERATURE REVIEW

SDMS Feasibility

Biopharmaceutical companies are under tremendous pressure to bring drugs to market quickly, with low overhead costs. Quality control laboratories in particular, provide the "last line of defense" for users of a manufactured drug, by verifying the strength, integrity, safety, purity, and quality of drugs before their distribution. Managing, analyzing and sharing information within departments, across departments and sites remains a big challenge for most pharmaceutical companies. P. Rees noted that nowhere is this more apparent than in sending data from laboratory instruments to LIMS or other corporate software [19]. This is because it is more than just transferring data, but managing it too, including security and regulatory compliance [19].

In an article on instrument integration in Pharmaceutical Quality Assurance (QA) laboratories, Reed-Jones, states that, "it is surprising that most laboratories deal manually with data flows between analytical instruments and systems like SDMS and electronic laboratory notebooks" [20]. In my opinion, that it is not a surprise, because many people fear the unknown. Many quality control laboratory managers - although they are intrigued with the idea - have stayed away from automating laboratory process workflows because they are fearful of losing data, not being able to access data due to IT networks being unavailable when needed, etc. These can all be defined by what is known as the return on investment (ROI). Although direct cost savings is important to consider, compliance is probably the overwhelming factor [20]. Scientific organizations are faced with the challenge of managing large amounts of diverse scientific data from multiple heterogeneous data sources in a compliant manner. This challenge along with the

increased migration towards a paperless laboratory and electronic laboratory notebooks offers promising opportunities for SDMS [21]. Integrating a SDMS with other instrumentation and applications in a quality control laboratory provides significant benefits (refer to section on the "Benefits of Integrating a SDMS (with laboratory instrumentation") both within the laboratory and the organization collectively.

Benefits of Integrating a SDMS (with laboratory instrumentation)

Integrating a SDMS with laboratory instrumentation not only can make it the central repository for data, but provides quality, compliance, costs and standardization benefits. As mentioned previously, quality control laboratories have a regulatory responsibility to extensively review all laboratory data for product strength, integrity, safety, purity, and quality. Transcription errors are a major source of potential errors. The use of reliable instrument interfacing provides a significant way to reduce human errors. It was indicated by Reed-Jones that "3% error in each level of transcription, which reduces to 0.5% with checking, is a generally accepted statistic" [20]. By eliminating the need for human transcription, this entire error source can be eliminated.

Although the full ROI may be hard to quantify, product recalls can be extremely expensive and may for some smaller companies be significantly damaging to their survival and existence. Statistics show that about 80% of pharmaceutical information is unstructured and not easily searchable. It can surely be inferred that laboratory data contributed in part to this percentage. Managers of quality control (QC) laboratories that still struggle with the idea of data integration and lab automation for the fear of having to protect their existing IT investments, may find that irrespective of the analytical instrumentation that are available, integrating this systems with a SDMS will take them

along the path of fully functional data integration, without having to abandon current instruments and other existing tools.

As of February 20, 2003, the Federal Drug Administration (FDA) announced that it has completed the first steps of its broad initiative to improve regulations for Pharmaceutical manufacturing which outlines the controls necessary to use electronic records and electronic signatures [22]. Without proper planning and interpretation, this can lead to cumbersome and excessive validation activities and over engineered solutions with no real benefit to the process. This is where the integration of instruments and systems like SDMS come into play.

With integration, the automated processes to collect and manage data do not permit actions to be missed, thus ensuring adherence to laboratory procedures. The greatest benefits [for cost reduction] will be seen in routine/high volume environments [20]. As is said when talking about LIMS, the level of benefit depends on current laboratory practices; therefore, there is not surprisingly a roughly linear relationship between the volume of results and time saved [23].

Laboratory automation through instrument interfacing can create a high degree of standardization across laboratories and the company. The benefits of this are increasingly being recognized and can be implemented on a global scale. Since interfacing can be implemented across remote sites, the global roll out of procedures, control, reporting, training etc. can be achieved from instrument to instrument and set up exactly from lab to lab [20]. Standardization in the way data is collected, treated and stored, results in an increase in lab productivity [20].

Instrument Integration Categorization

Instrument integration is more than just the mere physical connectivity of two or more instruments. According to Reed-Jones [20], instrument integration can be split into three categories: Basic, Standard and Advanced. Each category is briefly described below.

- Basic
    - Connectivity between instrument and target system for limited raw data results – in this case a SDMS
    - No application intelligence
    - No attempt to optimize or automate analytical testing
    - This level of integration is natively a part of some SDMS.
- Standard
    - Bi-directional integration – analysis requirements are downloaded to the instrument and instrument results are processed, transformed and enhanced prior to storage to the SDMS.
    - Eliminates manual operations by analysts to achieve the correct results for reporting.
- Advanced
    - Specialized algorithms and processes applied to support specific applications (i.e. handling unknown compounds by chromatography and content uniformity).

It can be interpreted by the descriptions above that the level of benefits for automation increase as you move up the continuum (Basic→Standard→Advanced) from manual systems.

Software Tools and Requirements for Integrating a SDMS

Closer linkage of instruments and data management software means smoother processes; however, ease of setup, use and maintainability still remain a concern. Configurability is one of the big trends seen among LIMS [and SDMS] companies, says Frost and Sullivan sector analyst Charanya Ramachandran [19]. Companies like Agilent, Waters and others have extended the capabilities of their instruments, chromatography data systems (CDS) and SDMS software with the aim of satisfying the demand for tighter integration and configurability by their customers.

Commercial interfacing software can accept data from a range of laboratory instruments and can be configured to perform a host of tasks, including making calculations, managing the maintenance calibration and validation of balances or other laboratory equipment without its own software, says D. Liptrott -marketing manager of Labtronics [19]. Due to the constant usage of such systems like SDMS and LIMS by a QC laboratory, these interfaces must be robust. To compensate for this, the interface can be embedded in the instrument software itself.

LimsLink, the most popular integration product made by Labtronics, has the capability to be embedded directly within some of the most widely used CDS and SDMS, such as Waters Empower CDS and NuGenesis SDMS, as well as Agilent's Cerity ECM – a SDMS product. Per Labtronics, LimsLink is the industry standard for instrument interfacing [24]. LimsLimk provides a validated bi-directional [advanced interface]

connection between a SDMS or CDS and any LIMS [25]. LimLinkCDS allows users to access a LIMS directly from menus within their CDS [19] to create work-lists, run their analysis, etc. LimsLinkECM, an advanced integration solution for Agilent's Cerity ECM, enables customers to easily connect to any LIMS. LimsLink is said to be able to capture, parse, reformat and report RS232 raw data from any lab instrument with RS232 output.

Another software tool for integrating RS232 instruments with SDMS is Nexxis. Per a Laboratorytalk news release and Waters Corporation, Nexxis, also from Labtronics, not only collects data, it also controls instruments, records the analyst's observations, comments and descriptions, and creates reports that can be sent directly to SDMS [26, 27]. The combination of NuGenesis SDMS and Nexxis, as with Agilent's Cerity and LimsLinkECM, creates a robust data management and archival system for RS232 and TCP/IP based instruments.

Current Understanding

The majority of pharmaceutical companies whose laboratories would benefit the most from purchasing, and integrating a SDMS with their current analytical instruments and other software applications, have yet to. Select Science noted that this may be due to the lack (until recent times) of suitable products [20]. That may have been true at one time, but I tend to agree with the latter part of that statement, which says, "…combined with traditional heavy reliance on paper systems with Pharmaceutical Quality Assurance" [or Quality Control]. This is very true for the laboratories in which I have worked over the last 11 years.

As it was noted by Select Science, the benefits [of integrating a laboratory to automated process] are clear. The overriding element is the internal intelligence of the instrument integration software. The software needs to be of comparable sophistication to the LIMS, CDS or other instrument data analysis suite (i.e. SDMS). At the same time, users need a flexible yet simple product that will standardize the laboratory, whilst improving quality and compliance, and reducing company costs.

Research Question

Which currently marketed SDMS is most suited for a pharmaceutical laboratory?

Intended Research Project

The intent of this thesis research project is to use various data collection methods to determine which currently marketed SDMS is most suited for a pharmaceutical laboratory based on the defined user requirements. This includes how these systems are also validated for use. My primary and secondary audiences will be scientists in research and quality control laboratories, respectively.

CHAPTER 3: METHODOLOGY

Participants

The participants in this thesis were selected based on the accessibility of information and limitations of company sensitive information. The SDMS evaluation focused on six systems, but the usability of them only focused on two laboratories. These two laboratories used different SDMS, and provided different applications for each.

Treatments

Representatives from each laboratory were given a questionnaire to complete in a two week period. Based on vendor audits and user requirements for the laboratories represented, the questionnaire results focused on two SDMS, which were Waters NuGenesis SDMS, and Agilent Cerity ECM. These systems provide different advantages that were evaluated and compared.

Procedures

The methodological approach used consists of the following steps:

Phase 1: Development of a User Requirements Specification document

Importance of User Requirements Gathering

Requirements' gathering is essential to the software development and implementation process. Requirements specifications documents vary based on the intended functionality of the system. When defining requirements you are defining exactly what the software must do, but not how it must be built. There are key functional

and non-functional requirements needed to completely develop a well-defined requirements specification document.

The functional requirements define the fundamental actions that must take place to ensure the system works as required by the customer or System Owner. These requirements may include, but are not limited to:

- Regulatory requirements;
- Business requirements;
- Interface requirements;
- Data requirements;
- Error handling;
- Reporting requirements;
- Performance requirements

The non-functional requirements specify criteria that can be used to judge the operation of a system, rather than specific behaviors. These requirements may include, but are not limited to:

- Accessibility
- Look and Feel
- Robustness
- Scalability
- Usability
- Platform compatibility
- Supportability

Format for Specifying User Requirements

Requirements' gathering is a unique process that consists of creating more than just a single document. It is a collection of specifications created with the user's input via interviews, brainstorming sessions, role plays, etc.

These specifications include:

- An introductory description of the project background and purpose for the system to be purchased or built.
- A diagram model that helps to set the context of the system to be built.
- A list of uniquely numbered statements with dependencies
- Use-case diagrams and descriptions using the Unified Modeling Language (UML) to capture the interface interactions between the user and the system. This puts the requirements in a form of interactions in a familiar context for the user, as requirement statements may not be intuitive to every user.

Selecting a Requirements Specification Template

Requirements' gathering is such a unique process depending of the specific needs of the client's use for a software application, that there is no one perfect method for gathering requirements. Whether you prefer a written document, screen diagrams, prototyping, or use cases, the most important outcome is that the people who need to understand the requirements can do so. The intent here is not to imply that all formats are identical, but if the user does not understand UML for example, they may not be able to identify any errors.

There are numerous templates available for gathering requirements; however, this project uses the Volere Requirements Specification template as this process offers a complete solution for gathering requirements in a way that does not become a project itself. The template is a great way to ensure that all relevant areas have been considered. The requirement shell, also called the "Atomic Requirements Template", is a convenient repository that ensures uniform and completely documented requirements. It is very specific and could help bridge communication gaps between users and developers if filled out completely and correctly.

The purpose of the requirements for this research is to define the capabilities and characteristics to be used in designing or evaluating designs for a SDMS.

The motivation is to give readers a basis for comparison of some of the most commonly used commercial SDMS software packages, and as an aid in selection for future reference by providing a baseline for both validation and verification. The list is far from all-inclusive. It is meant to include only SDMS software for pharmaceutical and biotech quality control and research laboratories. The programs included tend to be more common in terms of exposure, use and review. The SDMS programs researched were selected based on a review of the description of numerous data management systems on the market that claim to have the capability of being used in a scientific laboratory environment. This list was narrowed further based on vendor response and cooperation. All notations reflect the capabilities of the latest version of the software at the time of the comparison.

A list of requirements to consider is listed in Appendix B. For examples on how to go about trawling for requirements, refer to the book ‘Mastering the Requirements Process' [28] or the Web site - http://www.volere.co.uk.

Phase 2: General analysis of SDMS literature on the key functionality and technology.

1. Brochures
2. Compliance documents
3. Published works
4. Vendor demonstrations
5. Evaluation against requirements (see Appendix C)

Phase 3: Evaluation of Current Use in Pharmaceutical laboratories

1. Questionnaire (refer to the Results section and Appendix D)

CHAPTER 4: RESULTS

Overview

The results are reported in three main sections – Phases 2, 3, and 4, respectively. Phase 2 is a general analysis of SDMS six vendor software applications and comparison to the user requirements as defined in the Chapter 3 (Methods). Each SDMS application was analyzed by the methods described below:

1. NuGenesis SDMS
    a. Brochures
    b. Published technical and compliance documents
    c. Hands-on classroom experience
2. Cerity ECM
    a. Brochures
    b. Published technical and compliance documents
    c. Hands-on classroom experience
3. TargetWatch
    a. Brochures
    b. Published technical and compliance documents
4. E-Flexion
    a. Brochures
    b. Published technical and compliance documents
5. Biotrue CDMS
    a. Brochures

b. Published technical and compliance documents

c. Teleconference and Webex demo by vendor

6. Abrevity FDM

a. Brochures

b. Published technical and compliance documents

c. Hands-on experience via temporary online 30-day access

Phase 3 is the study of the current use of the SDMS software in pharmaceutical laboratories. No additional results are reported for Phase 1, as it was the development of the user requirements specification document, which was developed in April 2006. General results of the SDMS vendor review and software validation life cycles are discussed later in this chapter; however, a detailed comparison of how the functionalities of each SDMS compare to the user requirements are displayed in the matrix table in Appendix C.

Summary of findings

Phase 1: Development of a User Requirements Specification document

Figure 5 (below) is a graphical illustration of the number of requirements that were defined per category based on Volere Requirements Specification template. Refer to Appendix B for the entire Requirements Specification document.

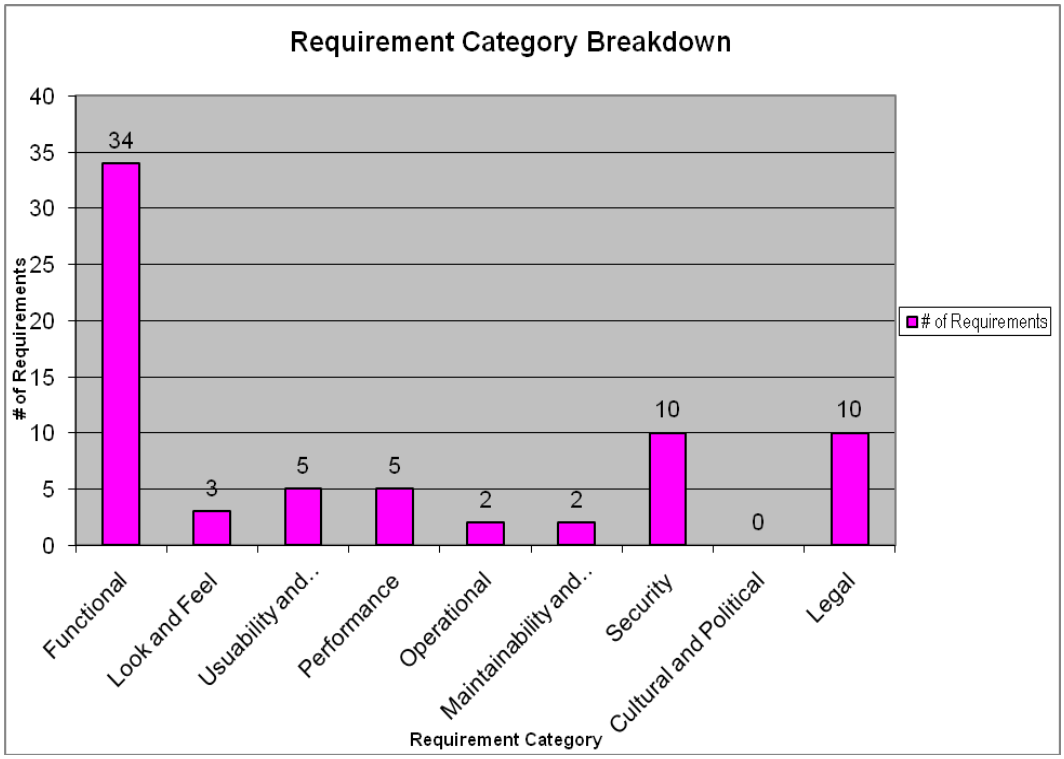


Figure 5: Requirements by Volere Category

Phase 2: Vendor Software Analysis

Phase 2 was designed to provide a general overview of the strengths and weaknesses of the SDMS software products that were analyzed and compared to the requirements.

NuGenesis SDMS (Waters Corporation)

The Waters NuGenesis Scientific Data Management System (SDMS) is the nerve center of Waters Laboratory Informatics' suite of software solutions [29].

This information management platform is an automated electronic repository that stores and manages all types of scientific data to a centralized database, offering excellent integration with the majority of the applications that researchers use. NuGenesis SDMS provides the foundation for scientific data preservation [29].

NuGenesis SDMS provides users with one unified platform for file and print data (see Figure 6). This technology makes it possible to view data from anywhere on a company's network. NuGenesis SDMS is built on Oracle 10g technology, and is scalable to a company's needs, by supporting data from instruments based on UNIX, MS Windows, and Mac OS. The administration/control of NuGenesis can be Web-based or distributed to local computers, depending on company needs.

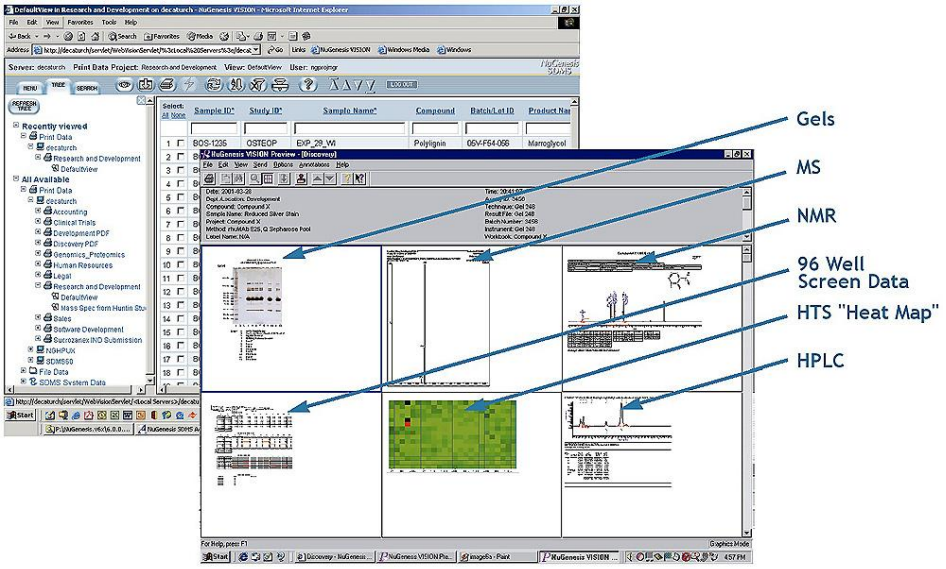


Figure 6: NuGenesis – Print Data organized by project

(Used with the permission of Waters Corporation)

NuGenesis manages print data from instruments such as HPLC and GC data reports (see Figure 7). It also supports raw instrument data such as spectra and chromatograms, chemical structures and reactions, spreadsheets, presentation, and other document files. NuGenesis is vendor-neutral and manages internationally standardized data exchange formats such as JCAMP-DX. NuGenesis SDMS also allows the user to revise previously captured data; however, to revise a record a new report ID is created,

for records with the same name. The user is allowed to enter an annotation to further distinguish between the records.

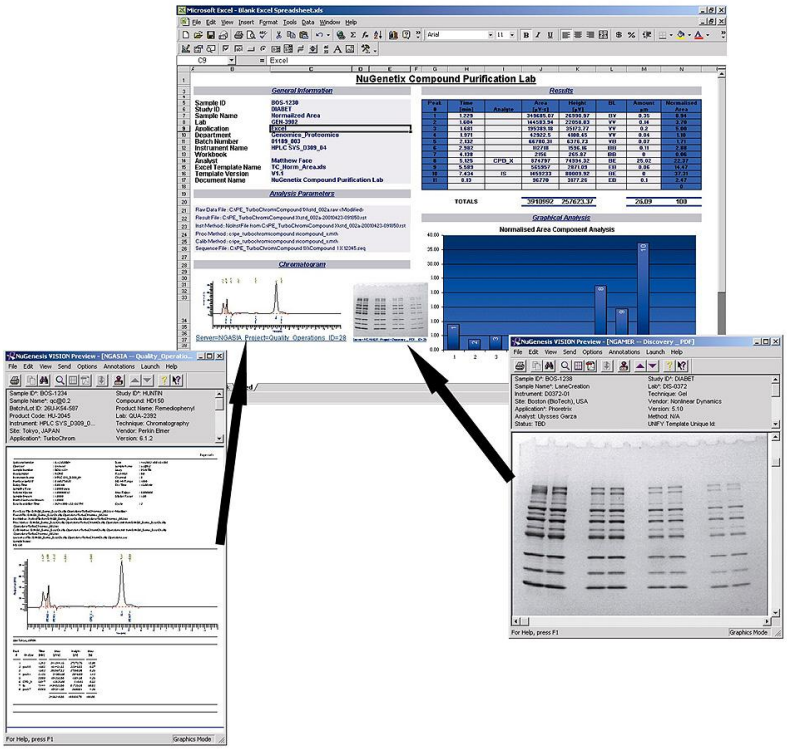


Figure 7: NuGenesis-Print Data from other Applications

(Used with the permission of Waters Corporation.)

NuGenesis has search filters and retrieval tools that allows a user to find text included in graphs, printed reports, tables, etc based on metadata tags. It also allows for simultaneous viewing of multiple reports from disparate sources. Data can be viewed in its standard format without launching the source application, and instrument data can easily be restored. NuGenesis also provides a means for scheduled archiving, based on a company's policies and requirements. The archive agent periodically scans the file system for new or changed files and automatically copies them into the SDMS database. At the same time, metadata for the information is extracted, cataloged, and stored in

context, making information easier to find. The NuGenesis SDMS platform also meets the federal regulations of 21 CFR Part 11 and intellectual property technical requirements for electronic records and electronic signatures.

Refer to the section entitled 'Phase 3: User Questionnaire' for some of the weaknesses of NuGenesis SDMS per current users.

Cerity Enterprise Content Manager (ECM) (Agilent Technologies)

Cerity Enterprise Content Manager is a software platform that provides a secure, central repository and rich content services to create, capture, manage, archive, and re-use business critical information scattered across the enterprise [30]. These records can include any type of electronic record – images, documents, presentations, and spreadsheets, or scientific information, such as raw data, SOP's and reports (see Figure 8).

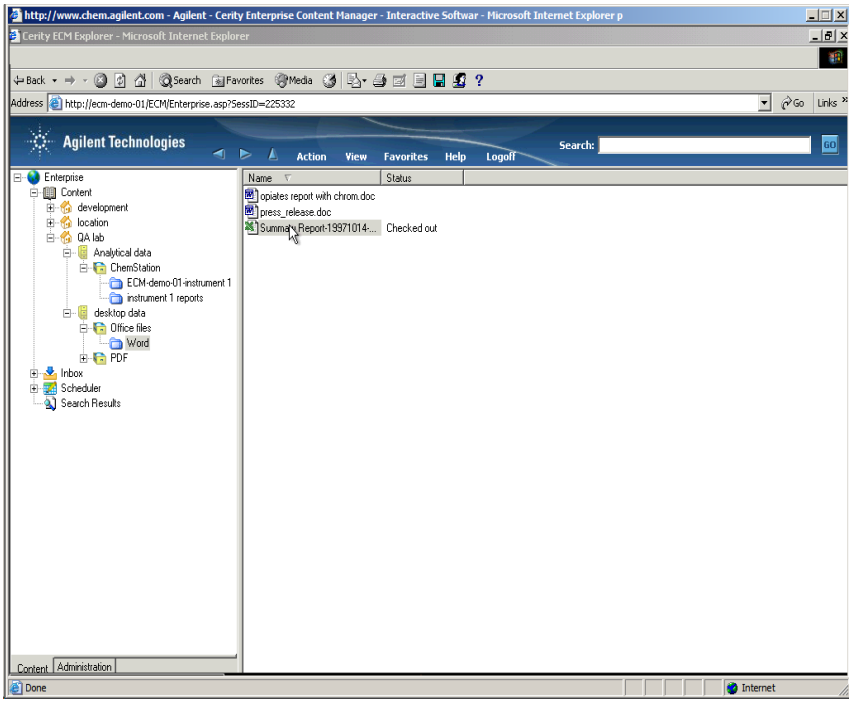


Figure 8: Cerity - Main screen and Structure Hierarchy

(Used with the permission of Agilent Corporation.)

Cerity ECM is capable of supporting all platform requirements as listed in the requirements specifications, including UNIX, MS Windows, etc. The Cerity platform is fully scalable from a single user, on an independent personal computer, to enterprise wide deployment – thus the name – Enterprise Content Manager.

Cerity ECM supports both raw and human-readable data formats, as well as metadata extraction and cataloging from analytical and all Word processing applications, and automatic archiving. Electronic records can be transferred into Cerity ECM from PC, Unix or Macintosh-based systems in a number of ways including print capture, uploading from third party applications, and manually via the Web Client interface, to name a few. Cerity ECM files can be stored on a protected hard drive or on a secure server. Files can also be archived and kept on-line or stored off-line using a management storage device application (see Figure 9). One advantage that Cerity ECM has over archival storage devices is that backup copies of its repository can easily be created by asynchronous mirroring, or replication to a remote site for disaster recovery. The files being stored in the application are automatically indexed and cataloged based on file properties contained in the application, such as, name, version, upload user and date, modified user and date, and so on (see Figure 10). When these files are revised, Cerity ECM creates a new version. Cerity ECM keeps a complete revision history of each file. At anytime, the current version of a file can be opened, viewed and compared to a previous version of the same file. This is in no way a complete list. An unlimited number of user-defined keys can also be assigned to a file using Cerity's Smart Filter Extraction feature using application plug-ins (see Figure 11).

Disk mirroring is the replication of logical disk volumes onto separate physical hard disks in real time to ensure continuous availability [31]. In addition to providing an additional copy of the data for the purpose of redundancy, disk mirroring – which is usually synchronous - can allow each disk to be accessed separately for reading purposes.

Synchronous mirroring of data is good for an internal data center, and externally for short distances; whereas, Asynchronous mirroring offers advantages when the mirror is located at a remote site. Some of these advantages include:

- No distance limitation;
- No performance degradation
- Scalable from small to large enterprise environments
- Significantly lower overall cost for remote operations
- Multiple remote servers can asynchronously mirror to a single data-center

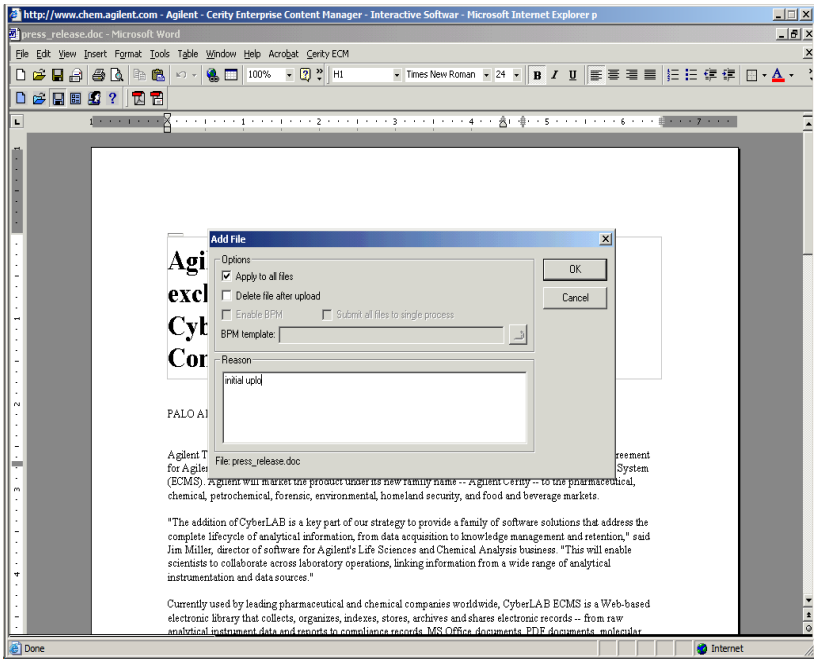


Figure 9: Cerity - File Save

(Used with the permission of Agilent Corporation.)

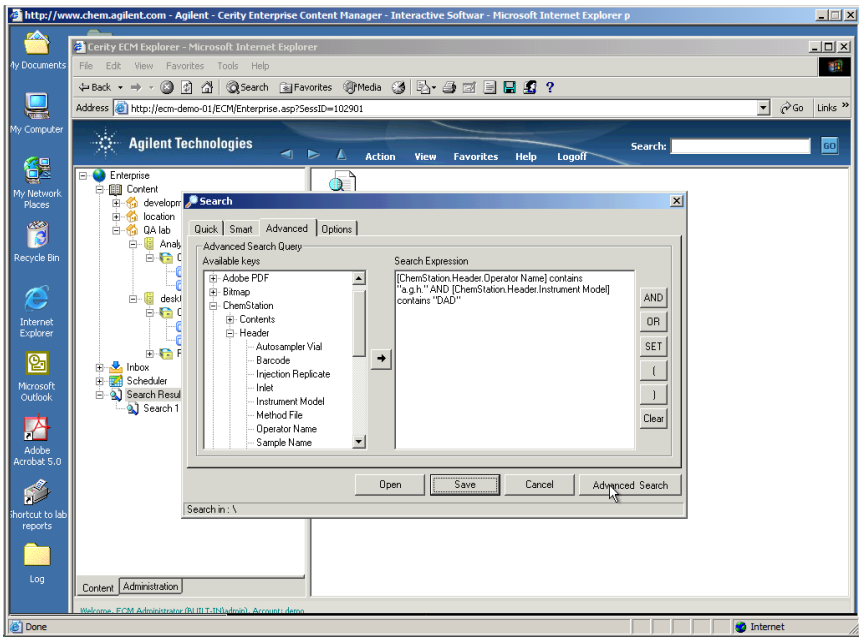


Figure 10: Cerity - Metadata Query

(Used with the permission of Agilent Corporation.)

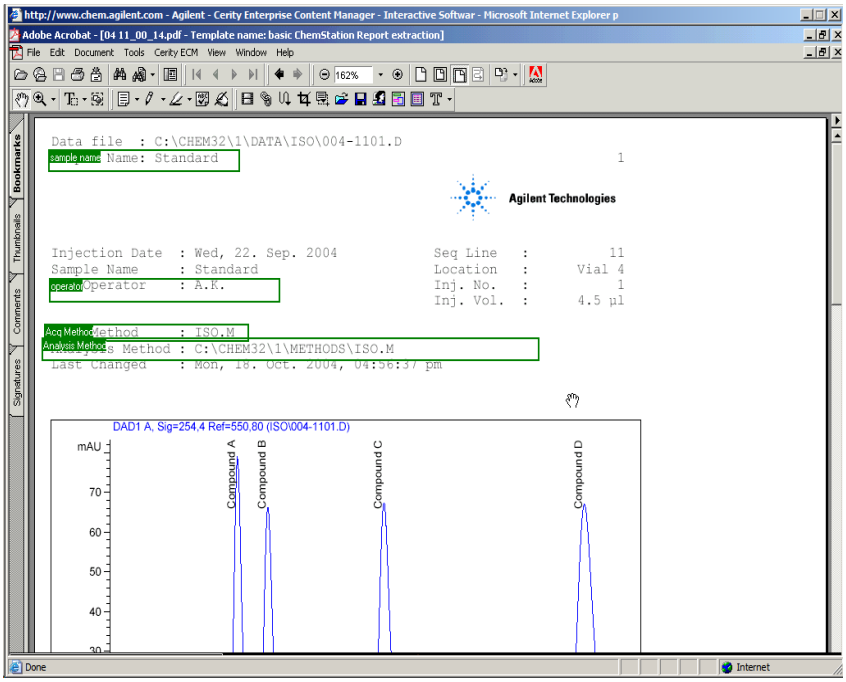


Figure 11: Cerity - Smart Filter Extraction (using plug-ins for other Applications)

(Used with the permission of Agilent Corporation.)

Cerity ECM addresses the main compliance record management issues (including data security) through encryption technology. Data integrity is maintained by versioning, and all main regulatory guidelines for 21 CFR Part 11, GMP, GLP and Sarbanes-Oxley (see Figures 12 and 13).

Refer to the section entitled ‘Phase 3: User Questionnaire’ for some of the weaknesses of Cerity ECM per current users.

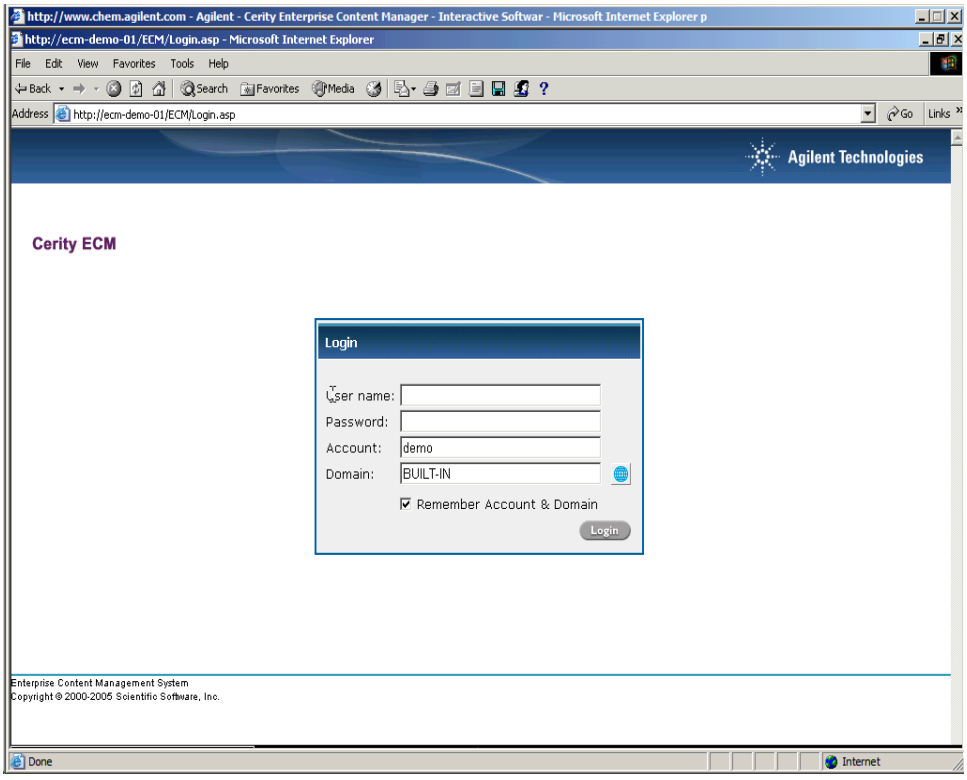


Figure 12: Cerity - Secure Login

(Used with the permission of Agilent Corporation.)

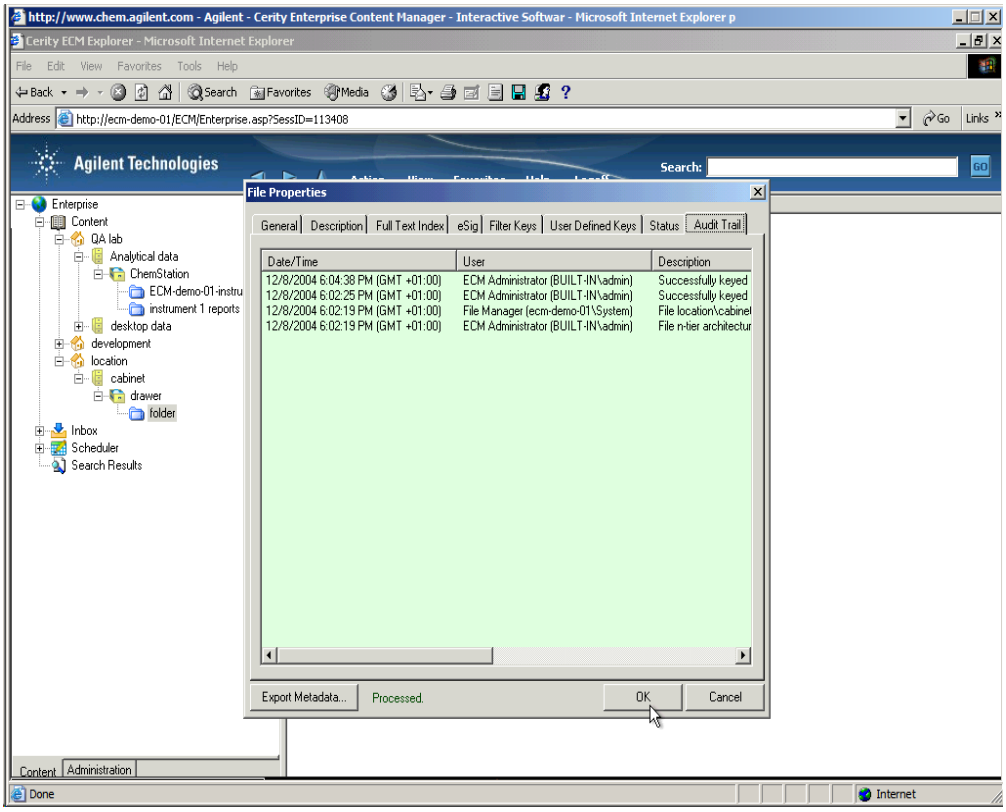


Figure 13: Cerity - Audit Trail

(Used with the permission of Agilent Corporation.)

TargetWatch™ (Amartus)

Amartus developed its TargetWatch™ SDMS solution to support the needs of scientific research team. Using the TargetWatch™ SDMS organizations can streamline the management of scientific research data. TargetWatch™ provides scientists with a single common integrated interface where they can access, store and organize all data relevant to research projects and share this with other team members [32] (see Figure 14).

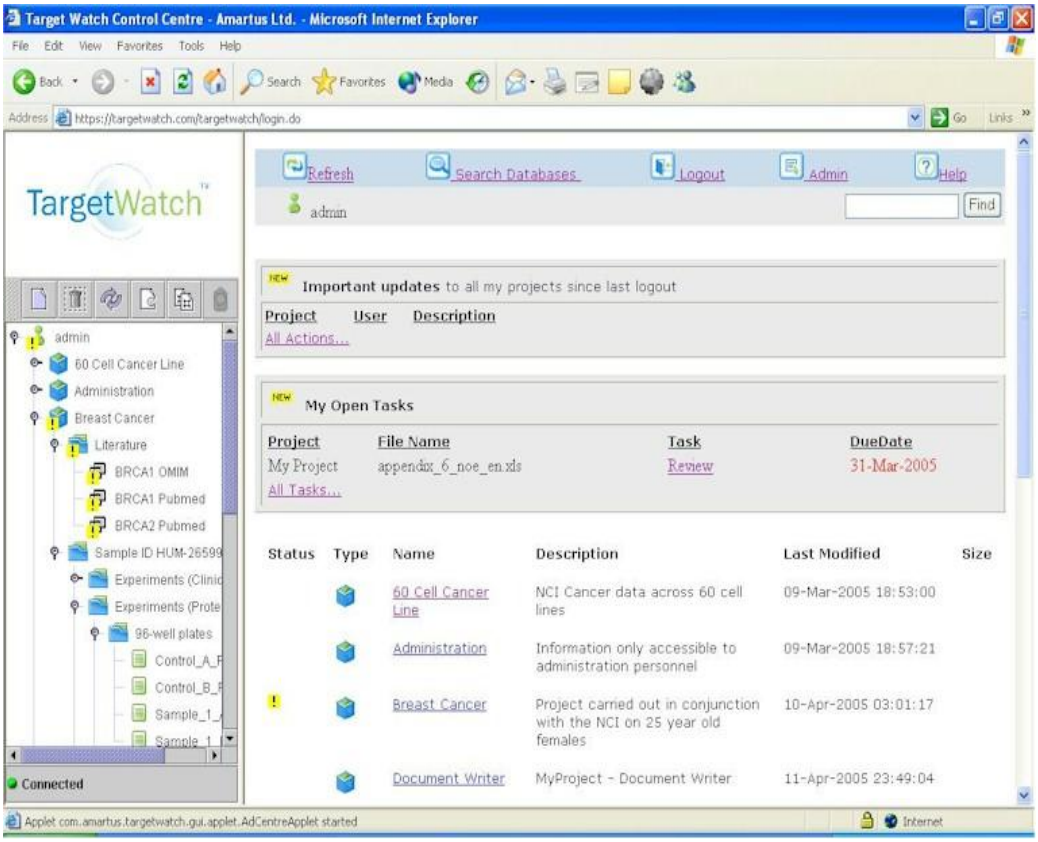


Figure 14: Target - Watch Main Screen

(Used with the permission of Amartus.)

Project areas hold all documents and data-files associated with a project. In addition researchers are able to search scientific databases and store important searches, results and annotations in projects for future references. This integrated data management approach enables scientist to truly capture all important data in a single place (see Figure 15).

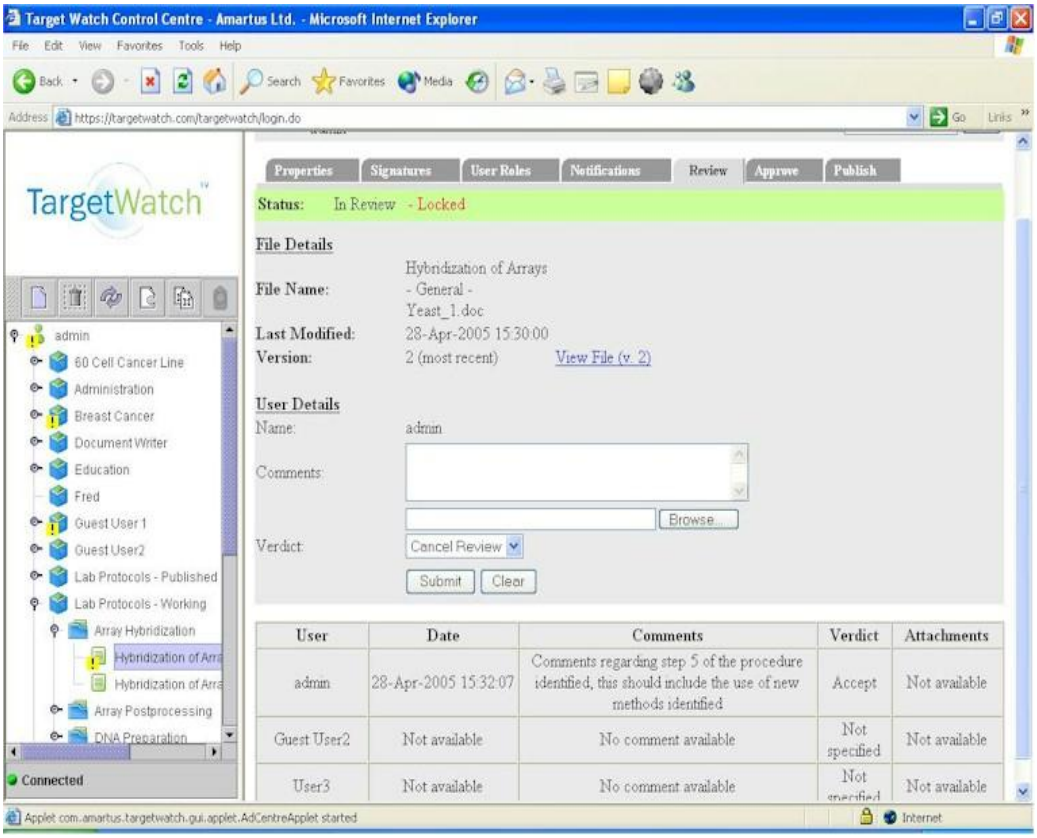


Figure 15: Target - Watch Project View

(Used with the permission of Amartus.)

TargetWatch™ provides advanced tools to manage and process data including:

•Integrated database access, search results capture, annotation, automated rerun and alerting all designed to assist with information processing in a dynamic environment.

•On-line document workflow designed to automate key signoff and approval.

•Version control, audit logging & electronic signature to support data integrity and traceability.

•Metadata and content searching to search and locate relevant data.

TargetWatch™ is built on Industry Standard JAVA/J2EE Application Server and Oracle® 10g relational database technologies, and has been thoroughly tested on all specified Windows platforms listed in the requirements specification. The company will be launching a release to run on Linux in the near future. TargetWatch™ has a scalable architecture, open standard interfaces for integration with existing IT infrastructure, comprehensive data security, and a simplified backup and disaster recovery option.

TargetWatch™ meets the federal regulations of 21 CFR Part 11 and intellectual property technical requirements for electronic records and electronic signatures. A matrix of how TargetWatch addresses ER/ES requirements was made available by the vendor and reviewed as part of this analysis. In addition to the IP requirement for which TargetWatch™ was designed, it has also been designed as an integration platform providing access to all of the scientific data sources required by researchers on a daily basis.

TargetWatch™ can manage files generated from any instrument or a desktop application. The Import/Export Manager component provides a programmable interface through which scheduled or on demand upload of data can be achieved. In addition users can manually Import / Export data files from any location that is accessible to them directly from a Web browser. For more advanced integration covering specific complex instrument integration, Amartus is currently looking at integrating third party solutions such as Labtronics & Csols who support a wide range of lab instruments.

TargetWatch™ supports automatic archiving and retention, and the use of metadata tags for cataloging its stored data and information. It is fully audit trailed, and has full-library services including check-in, check-out and revision history.

The one disadvantage to TargetWatch™ as it relates to the scope of this project is that it was primarily designed to support groups in early stage research (pre-clinical).

E-Flexion (Computer Compliance, Inc.)

E-Flexion™ is an Automated Data Management System (ADMS), which benefits each division of the pharmaceutical/biotech or Medical Device manufacturing industry [33]. E-Flexion's™ web and data repository and application server are built on MS Windows NT or Windows 2000 platforms. It could not be directly determined whether the other platform requirements for use with instruments based on UNIX etc can be supported. However, it can be inferred via the following summary that the application probably does support most disparate sources.

E-Flexion's™ core product currently has two primary components that are required at each plant site to run the application: The Process Scheduler, and the Web Portal. The Process Scheduler provides continual search, collection, and cataloging of data. It also supervises data analysis for processes configured to analyze data. The Web Portal provides authorized users from across the company's network a secure viewer into the detailed and summarized data for a site, process, or piece of equipment or production run. It also provides statistical charting and trending of data, as well as the ability to e-sign reports, and associate metadata files with a run. The EFlexion™ Directory [within the Web Portal] provides quick, one-click access to specific functions of EFlexion™. These functions are Overview, Systems, View the Data, Audit Log and Associate Files [32] (see Figure 16).

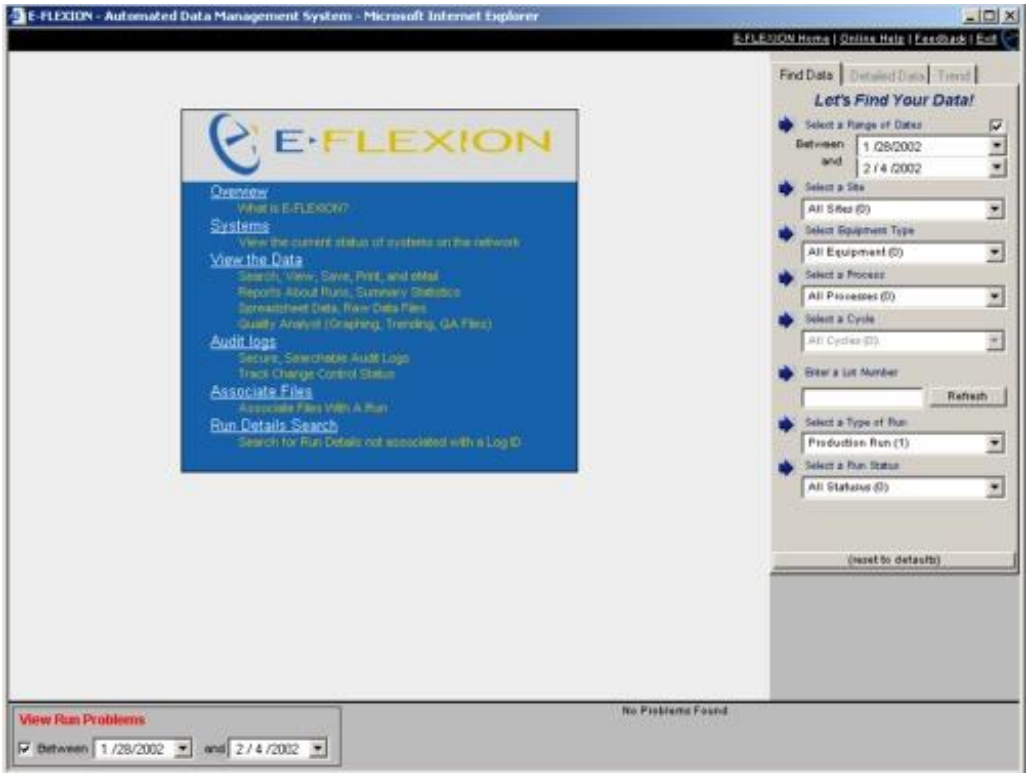


Figure 16: E-Flexion - Directory View

(Used with the permission of Computer Compliance, Inc.)

The Run Summary Data screen displays the results of the user's query. The results are color coded based on the run status [32] (see Figure 17).

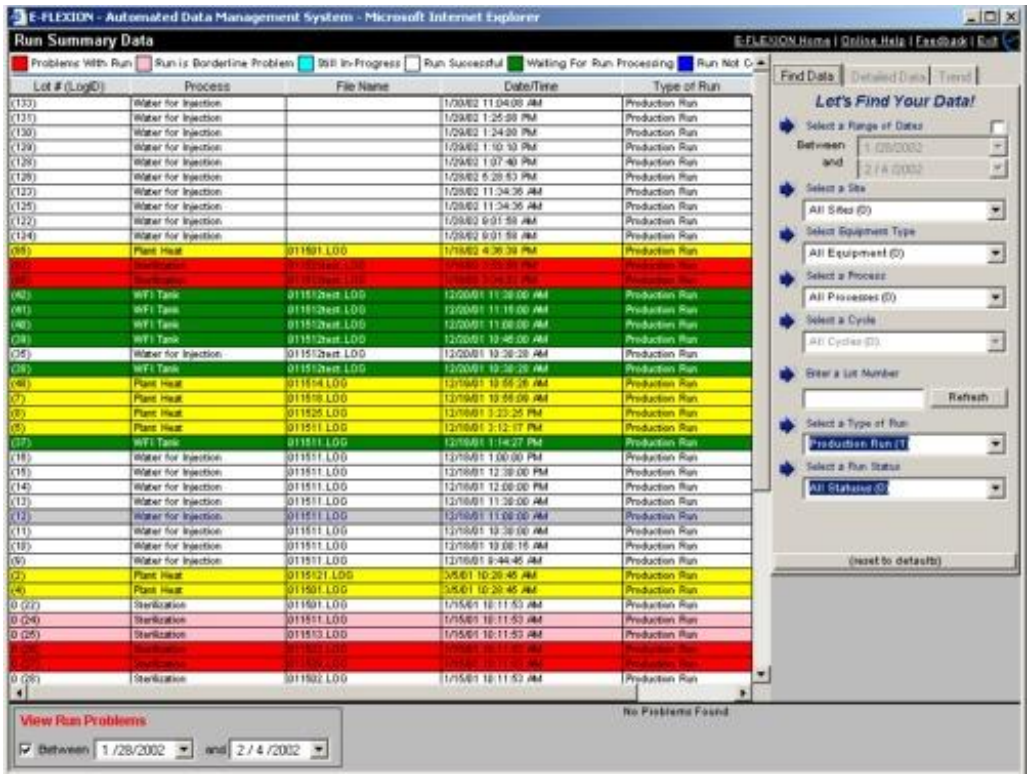


Figure 17: E-Flexion - Run Summary Query

(Used with the permission of Computer Compliance, Inc.)

The Written Report allows the user to view, print, save, and email the results of the analyzation process. Reports generated by E-Flexion can be electronically signed in compliance with 21 CFR Part 11 [32] (see Figure 18).

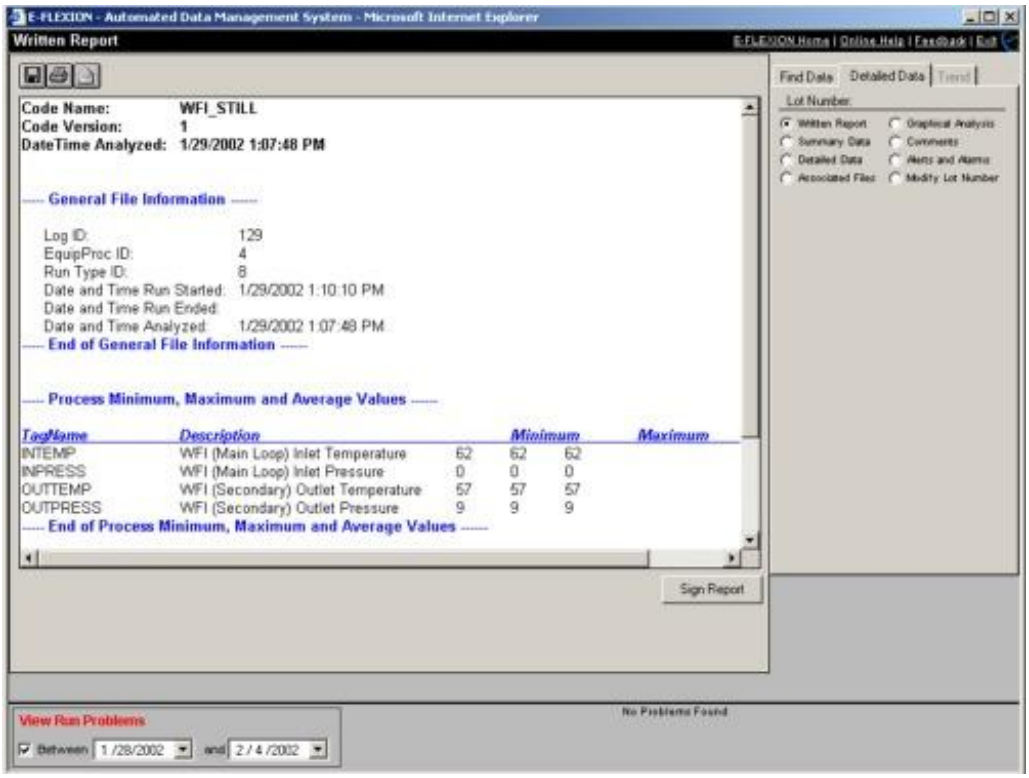


Figure 18: E-Flexion - Written Report

(Used with the permission of Computer Compliance, Inc.)

From the Graphical Analysis of Data option, users can plot statistical charts. Chart types include: Xbar, Range, Standard Deviation, Individuals, Process Capability Histogram, Run, and Scatter Diagram [32] (see Figure 19).

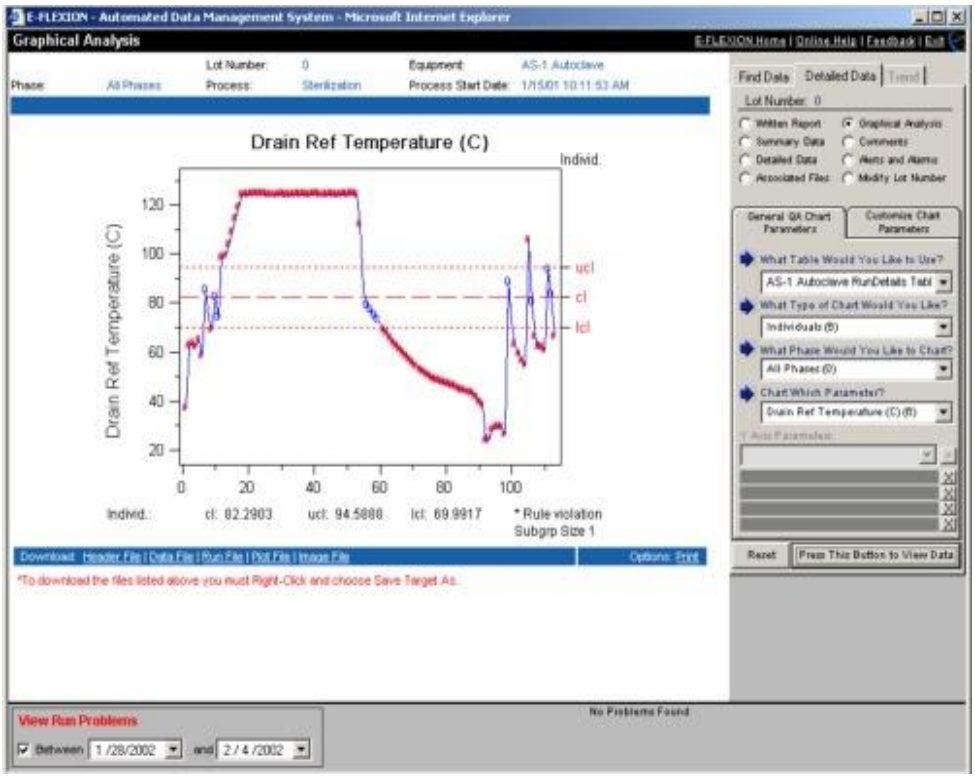


Figure 19: E-Flexion - Graphical Analysis Display

(Used with the permission of Computer Compliance, Inc.)

There are other secondary components, such as the Universal Analysis and Reporting and Universal Parsing modules, which are not discussed in this thesis. More information about these and other modules can be found on the Computer Compliance, Inc. Web site.

E-Flexion™ automates every step in its management of information, thus eliminating manual data handling and saving time for a user. With E-Flexion™, data is collected and analyzed 24 hours a day from any piece of equipment, for any analysis [32].

The system also meets the requirements of 21 CFR Part 11, having such features as an internal audit log (viewable only by authorized users) (see Figure 20) to track changes made any records in the database repository, The repository data cannot be deleted, allowing data transfer to the repository to be error-free. It also captures user ID and password based electronic signatures from authorized workstations.

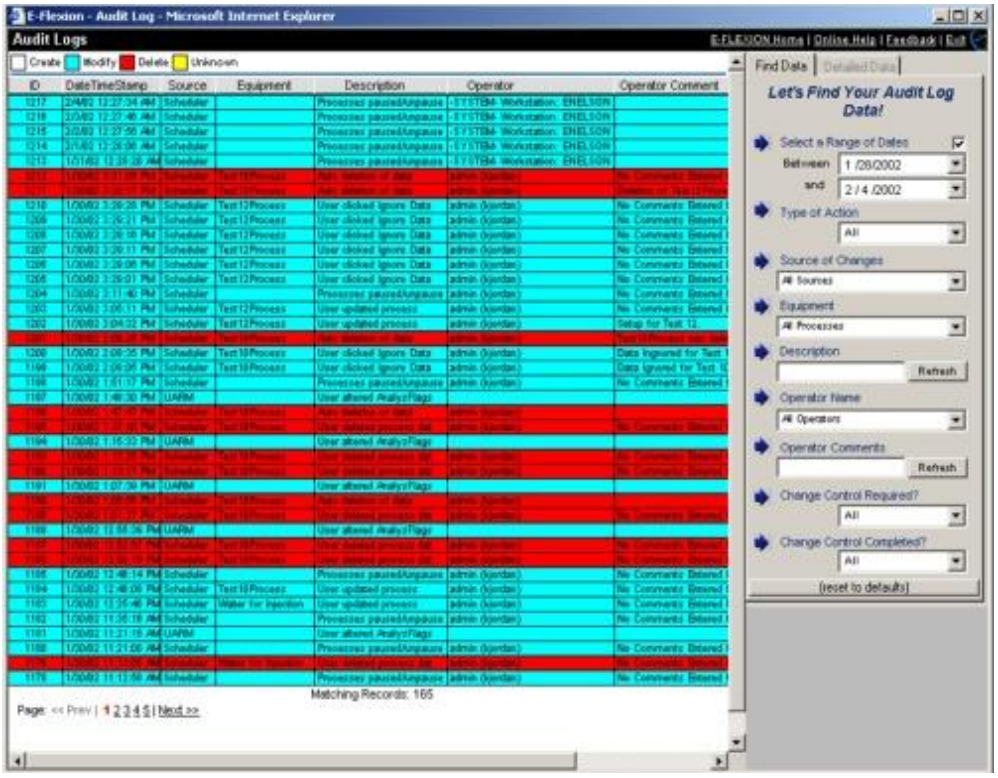


Figure 20: E-Flexion - Audit Log

(Used with the permission of Computer Compliance, Inc.)

E-Flexion's™ ADMS also has other functions related to process data archiving, manipulation, notification, and access. The system automatically archives data after each run or once a day to a secure network area on or off-site. The system integrates with stand-alone "file based" equipment, and automatically analyzes and summarizes raw

process data. Various system or process-related problems are made known to a user via email, which can include a copy of raw data to assist in understanding and resolving the issue. Authorized users also have the capability of associating ad-hoc files (i.e. custom reports) with a particular run. Finally, E-Flexion™ is capable of allowing comparative analysis of critical parameters for like-product and system performance trending.

E-Flexion does much more than collecting and archiving data. The data is analyzed and reported to reduce man-hours and increase time available for critical decisions and process improvement. The benefits of E-Flexion will have a positive impact on productivity and provide information for process improvement in addition to meeting Part 11 regulatory requirements.

FileData Manager (Abrevity)

ABREVITY's FileData Manager™ is the first low-cost software to transcend the limitations of traditional Information Lifecycle Management (ILM) solutions. Installed in minutes, easy-to-use and low-cost, FileData Manager's software empowers ILM via simple, yet powerful Information Discovery, Classification and Management technology. FileData Manager is compatible with CIFS and NFS and requires no server agents. A separate lightweight utility scans Windows and UNIX network or desktop storage systems and extracts target information. User can then quickly find and extract target words, values or phrases found inside file paths and common file types such as Microsoft Office, PDF, PST, txt, XML, HTML, etc [34].

Abrevity runs on any standard Windows platform, but can scan any CIFS or NFS shares. As far as its ability to collect and manage native instrument data, Abrevity can parse and extract all file path metadata, and extract values from inside .lei (Leica), fcs

(FACS 2.0 or 3.0), .csv or most common file types (.txt, MS Docs, PDF, PST, etc). Per the vendor, the application also supports other instrument file types as required and for a reasonable cost for the additional, required professional services. Abrevity interfaces with any standard CIFS or NFS file systems. If the lab instrument interfaces with a common network, it can generally scan the storage system

Abrevity provides software that empowers discovery, classification & management of files that can ensure compliance with 21CFR Part 11 requirements. It can also ingest compliance-related taxonomies to allow searching on those words and phrases. Full compliance will require other hardware, etc. (perhaps a WORM storage system). Abrevity scans only those volumes that IT provides access to, so the software does not interfere with any security measures in place. Abrevity's database captures file metadata only and is not involved with electronic record signing, which is not compliant with 21 CFR Part 11 rules 11.50(a), (b), 11.70, 11.100 and 11.200(a). It also does not meet rule 11.300, because it is said to be a software only solution. Vendor FDA Compliance documentation states that it is the responsibility of the customer's organization to maintain password authentication to its hardware systems.
Abrevity was found not to be user friendly. It was difficult to understand and navigate.

Collaborative Data Management System (CDMS) (Biotrue)

Biotrue's CDMS is a software system for biomedical research laboratories that enables scientists to easily store, manage and share all types of instrument and analytical data files [35]. Researchers using Biotrue's CDMS can store and manage a wide variety of instrument data types using an intuitive graphical interface (see Figure 21). With a simple click, you can easily view, manage, and retrieve your files according to key metadata. Intelligent thumbnails and other display features allow one to easily search for and retrieve files.

Using familiar directories of folders and documents, the CDMS allows one to easily manage data through a web browser. Biotrue CDMS will work on any MS Windows, Mac OS or Linux workstation that was made in the 21st century and is connected to the Internet. The software is considered easy to use if one has experience with web-based email or photo-sharing sites, but additional training may be needed for those that do not.

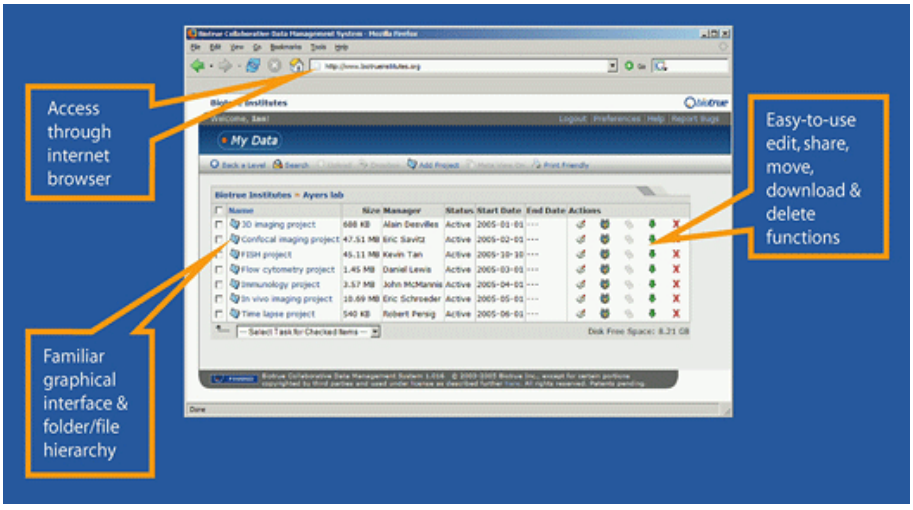


Figure 21: Biotrue - Graphical User Interface (GUI)

(*Used with the permission of Biotrue Inc.*)

The Biotrue CDMS manages multiple data types, including confocal microscopy, flow cytometry and office documents (see Figure 22). For certain data types, metadata such as reagents, instrument settings and other information related to data acquisition are parsed into a searchable database. For most images, thumbnails are created so one can quickly scan for files visually [35], instead of trying to find a long file name that was saved.

Files in Biotrue can be searched using metadata that was parsed when the data files were loaded into the CDMS (see Figure 23). It can also be used to locate files based on instrument settings across multiple data types, and from additional user-defined metadata (i.e. annotations).

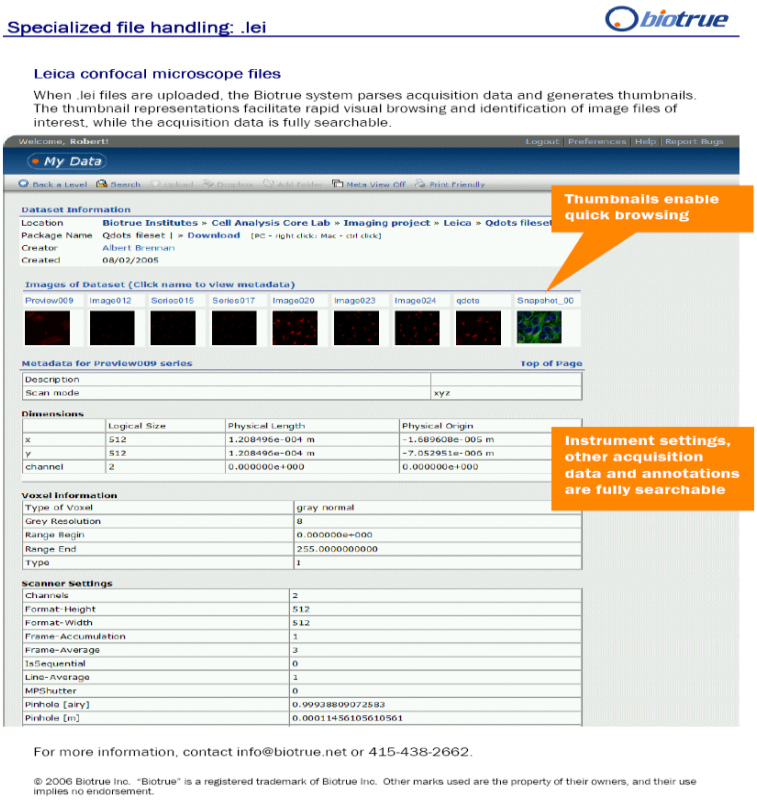


Figure 22: Biotrue - Datatype/Dataset Management

(Used with the permission of Biotrue Inc.)

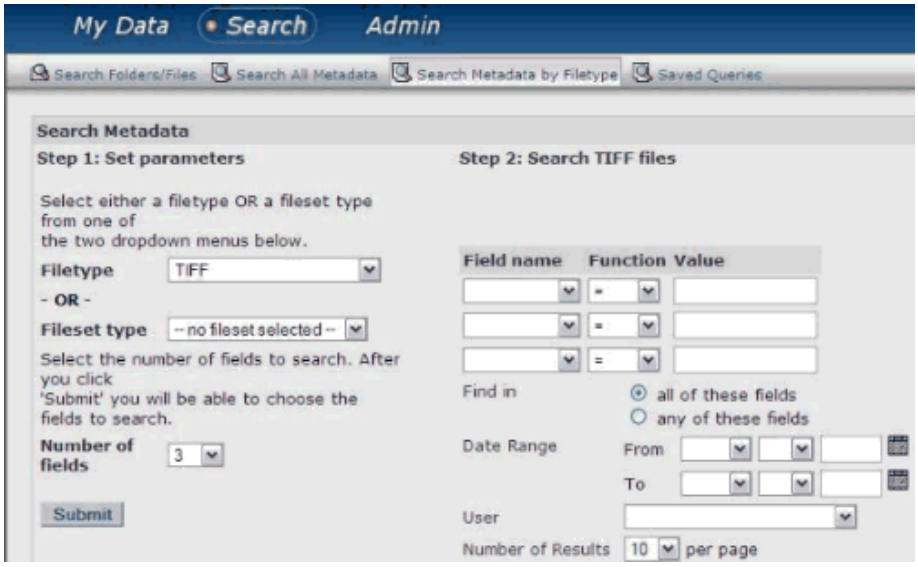


Figure 23: Biotrue - Metadata Searching

(Used with the permission of Biotrue Inc.)

The Biotrue CDMS is a secure database. Different users have different rights, and users with the right privileges can control access to data – per user, per folder [35] (see Figure 24).

Figure 24: Biotrue - Data Security

(*Used with the permission of Biotrue Inc.*)

Effective data management involves integration with other applications used for analysis within an organization. Biotrue has a published an application programming interface (API) that allows direct integration of other applications with its CDMS. Examples of this include the commercially marketed product Flowjo® software for analysis of flow cytometry data (see Figure 25).

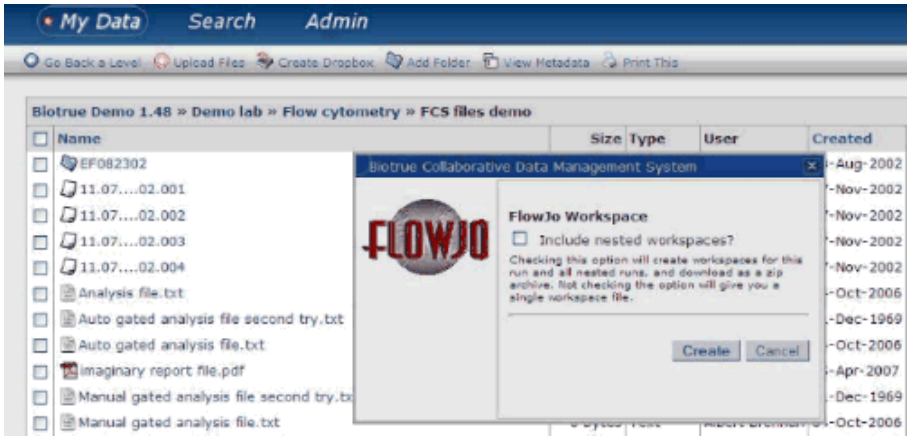


Figure 25: Biotrue – Application Integration

(Used with the permission of Biotrue Inc.)

A major advantage to using Biotrue CDMS is that data can be dragged–and-dropped into shared, permission-controlled storage (see Figure 26). Not all data management systems support this feature.

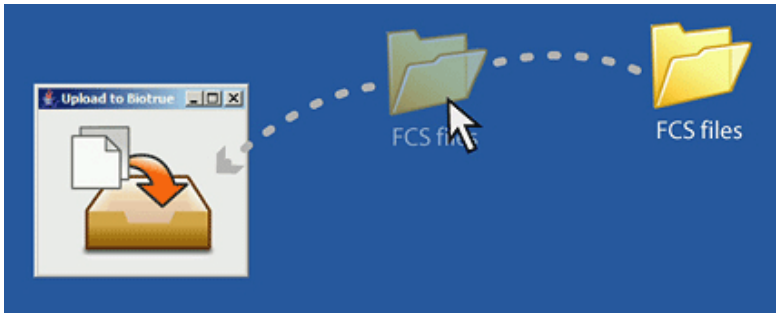


Figure 26: Biotrue - Drag-and-Drop

(Used with the permission of Biotrue Inc.)

A comparison to all the requirements sections could not be done for this product, as the vendor did not provide sufficient information to do so. The most significant functional requirement for which comparison data is missing is data archiving.

One disadvantage to Biotrue CMDS, besides any that may have already been mentioned, is that it was specifically designed to support biomedical research laboratories and not QC laboratories.

From the requirements that were gathered and the comparison that was done, NuGenesis SDMS and Cerity ECM stand out as two of the top applications available for use in a pharmaceutical QC laboratory. TargetWatch ™, a very competitive product, comes in a strong third, but as stated previously, it was designed to specifically meet the needs of pre-clinical laboratories (see Figures 27 and 28). This could pose some concerns and limitations for use in a GMP laboratory.

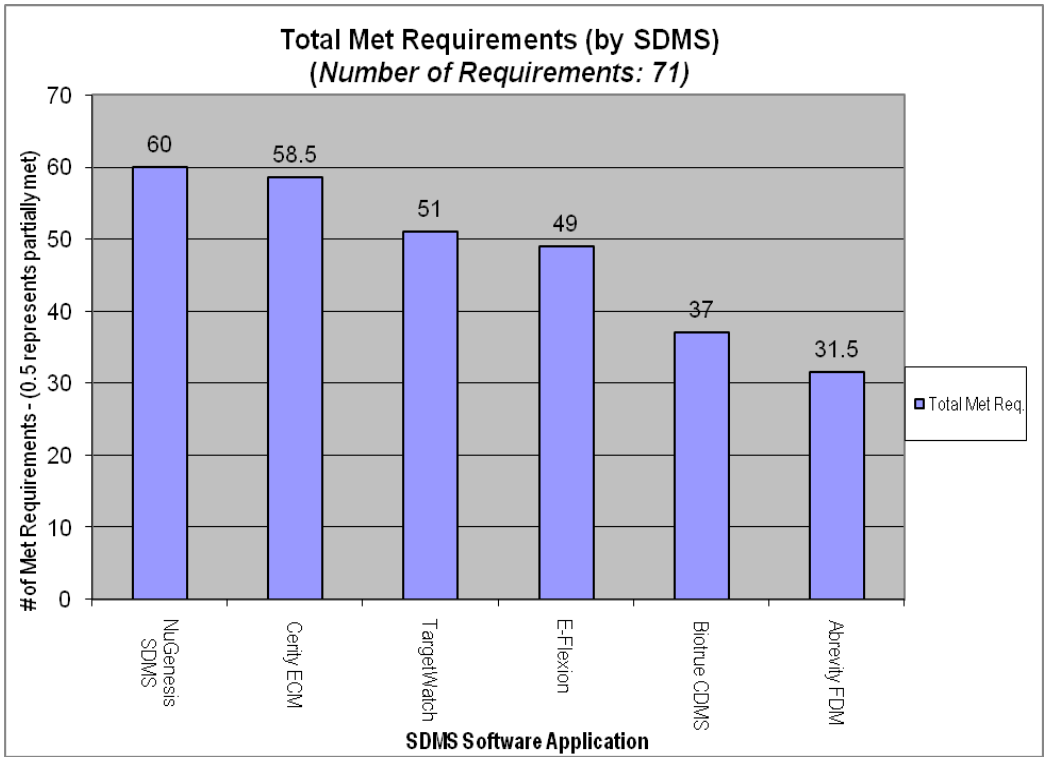


Figure 27: Total Met Requirements

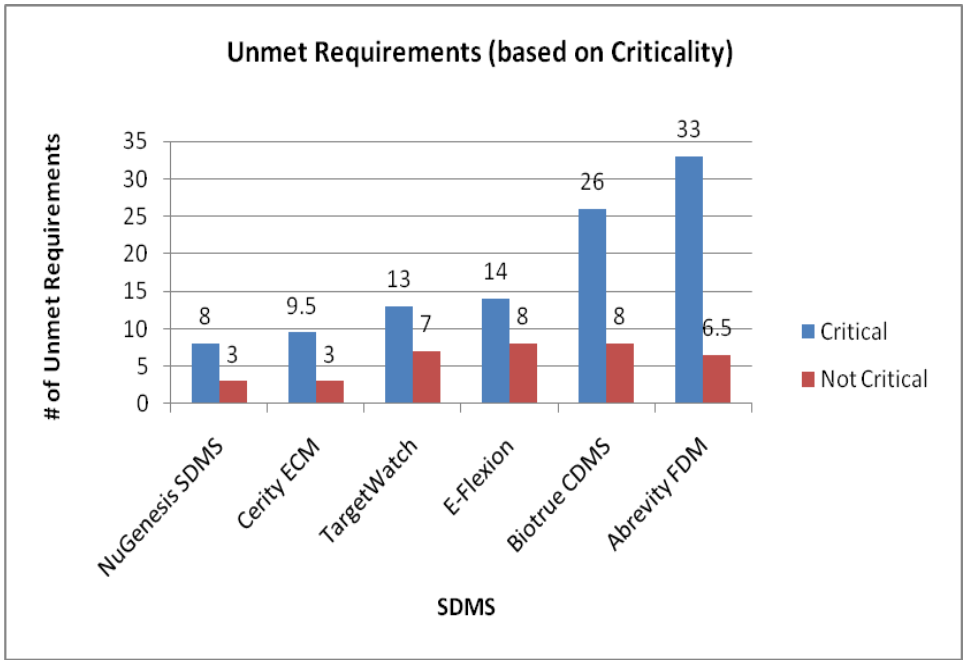


Figure 28: Total Unmet Requirements (based on criticality)

Phase 3: User Questionnaire

The purpose of this questionnaire was to identify which SDMS is being used in pharmaceutical laboratories, and gain general knowledge of how it is being used. The questionnaire was originally administered via email; however, the formatted questionnaire and specific results can be seen in Appendix D.

There are currently two different SDMS software applications being used at my current place of employment.

One area is using Agilent's Cerity ECM application as part of a larger global toolkit project and application. The Toolkit project was formed to consolidate over a dozen legacy systems (globally) providing one global system for Analytical Chemistry information about samples in the global Discovery organization, enable new business processes, and provide professional cradle-to-grave management of all related electronic data. (e.g. raw instrument data, instrument reports, instrument methods, etc.) [36].

This area had several key functionalities they required which other SDMS applications, such as NuGenesis did not provide at the time of purchase. At a high level these include, but are not limited to:

- Ability to keep files at each of the four global sites while still having only one database (federated storage)
- Support of real-time operational systems integration for files (and print capture).
- Ability to load sample on instrument, go back to desk and later be able to view files stored in Cerity, "through other external systems".

- Cerity generates XMLs with all metadata in them which we feed to external systems for automated integration.

Some of the minor issues included budget constraints, the fact that NuGenesis files are stored inside the database as binary language objects (BLOBs). Database disks are often very expensive. Cerity stores the file on simple windows file server (can be at any price level). Cerity has built in direct use of EMC Centera's (CAS device), and Tivoli. NuGenesis can only use them if one buys and put in place a third party application that makes them look like a file share. NuGenesis design and use has focused on doing print capture, using NuGenesis out of the box (very light on integration with external systems), and to long term archive files for GxP purposes. The Discovery laboratory tried to make NuGenesis work for their area to avoid using a different product than the other area in the company using a SDMS, but NuGenesis' focus is and has been for GxP areas which have very different requirements than Discovery.

The other area is using NuGenesis SDMS 7.0 SR-1 to improve data integrity and data security of several targeted systems across the department by providing compliant electronic record storage and audit trailing capability.

Some of the reasons for choosing NuGenesis SDMS were:

- The previous experience (of another colleague) showed good productivity increases.
- It could be implemented as a solution to address quality concerns of the lack of control over raw data in the laboratories.

When combined with the front-end security provided through qualified workstations, NuGenesis SDMS is expected to appreciably reduce security and data

integrity risks for the affected instruments. Waters Corporation offers software development kits (SDK) that provide tools to create custom workflows or data mining programs for automating processes for exchanging data between NuGenesis SDMS and other systems. In addition, operations in NuGenesis SDMS can be scripted to initiate activities in the systems it is integrated with, thus facilitating workflow processes and improving the exchange of data.

At the time of their purchase, little research was carried out on ThermoElectron's e-record manager and several other specialty niche products. This included Cerity ECM, which was still known as Cyberlab (and owned by Scientific Software Inc.) at that time. Agilent gained exclusive rights to market and sell Cyberlab in July 2004. None of them could stack up to NuGenesis SDMS, which had 70+% of the market at that time.

All areas agreed that the SDMS provided an improved efficiency, by allowing for a ‘paperless’ lab workflow environment, which reduces long term costs for archival management of paper artifacts, and allows for quick retrieval of relevant files or printed data from a central repository. More specific reasons are detailed in Appendix D.

CHAPTER 5: DISCUSSION

As pharmaceutical laboratories QC and R&D evolve from paper-based operations to a more modern electronic workflow to improve efficiencies, enhance compliance, and reduce time to market new compounds, there is a glaring need to manage the resulting data.

Although there are many types of systems available today that attempt to address data and record management issues, not all of them may be sufficient for a pharmaceutical laboratory's needs. Some of these are knowledge engineering/management systems, document management systems, content management systems, scientific data management systems, data archival systems and hierarchical storage management systems [2]. It can be inferred by these names that each one of these systems varies in its focus and capabilities. With the many different capabilities of each of these systems, it is extremely critical that one understands their current and future business work context, infrastructure, any regulatory requirements, budget constraints, and all other functional and non-functional requirements, so that a vendor does not sell them a standard interface that may not meet their needs. Of these different systems, a SDMS has proven to be very beneficial to the laboratories studied as part of this research project. After an in-depth review and analysis of the various data management systems discussed in Chapter 4 (Results), and the results of end-user questionnaires, it has been concluded that Waters NuGenesis SDMS provides the best ROI for the Quality Control laboratories; whereas, Agilent Cerity ECM was better suited (at the time of purchase) for the Discovery laboratory. This was mainly due to its inexpensive file storage and real-

time operational system integration. Files for Agilent Cerity ECM are stored on a file server; whereas, files for NuGenesis SDMS are stored in the database requiring larger and multiple databases. The Discovery laboratory was limited by budget constraints, which also contributed to their decision to go with Agilent Cerity.

This study showed that Waters NuGenesis SDMS provides three major benefits in terms of ROI. First is its usability – in terms of the ability to visualize information from various disparate systems within the appropriate context and its original format. Secondly is its scalability – as it allows for enterprise-wide scalability to cover data accessibility, reuse, and visualization requirements. Several other SDMS are built without scalability in mind and have a less than desired visualization appeal at their user interface. Lastly is the metadata aspect. The success of a pharmaceutical company is predicated on the transition of data throughout different stages of drug discovery, development, and manufacturing. NuGenesis SDMS provides such capabilities by enhancing the 'metadata' core environment for data searching and retrieval.

For any software integration project to be successful and compliant with the requirements and/or guidelines of the governing regulatory agency (i.e. the Food and Drug Administration for pharmaceutical companies), the software must be validated. The purchaser must perform a risk analysis and evaluation based on the intended use of the software by focusing on those functionalities with the highest impact on both the business and compliance requirements.

Consideration of findings in context of current knowledge

The rapid growth of scientific data in pharmaceutical laboratories, in addition to the need to protect intellectual property, is more prevalent today than ever before.

Though data management systems offer many benefits, many pharmaceutical companies have yet to link their analytical instruments to them. Until recently, it may have been due to the lack of suitable products, but it seems now that it is more of just the pharmaceutical laboratory tradition of heavy reliance on paper systems.

New analytical technologies, reporting requirements and regulations have forced a dramatic increase in the amount of unstructured electronic records such as instrument and image data files to reports. These file types have no common format between them. Consequently, this poses a challenge for laboratories of not only collecting these records, but managing them for long-term retention. Thus, SDMS play a critical role in making this data more manageable. Because NuGenesis SDMS serves as a central repository for such analytical data, it also serves as a management tool with respect to intellectual property archiving. With automated date and time stamping and electronic signatures and audit trails, the timeline for compound discovery will be clearly defined and defendable in a court of law should the need arise.

CHAPTER 6: CONCLUSION

Limitations of the Study

The primary limitations of this research project were first the necessity to restrict the number of scientific data management systems selected for this research. The ones selected tend to be more common in terms of exposure, use and review.

The second limitation was the number of laboratories that were currently using a Scientific Data Management System, combined with the lack of participation from other companies possibly using a SDMS.

Lastly, it was low enrollment in the INFO I-512 class, making the inter-rater reliability study – as outlined in the thesis proposal- an invalid research method.

Recommendations for Future Research

A number of enhancements and improvements for continued development of currently marketed scientific data management systems are in progress, as well as the development of new systems. Additional research possibilities include broadening the scope of this research to perform a detailed analysis of the requirements for a medical research laboratory to determine the feasibility of using one of the six aforementioned scientific data management systems for managing the laboratory's data and workflow.

Additionally, there are a few open-source application servers that have been developed for Experiment Management Systems (used in medical research laboratories) to allow an user to design his/her own schema online using forms in a Web browser, resulting in an easier navigation interface and allowing changes to the web forms to disseminate existing data. Considering the exponential increase in the volume of data,

combined with an increase of heterogeneous formats and autonomous systems, one may want to research the benefit of using or developing such an open-source application for scientific data management systems, which would allow for more flexible and powerful systems for research and quality control scientists. Finally, one may want to determine the possibility of creating a customizable user interface that would allow a user to customize the presentation of his/her data.

REFERENCES

[1] Jakobovits, R., Soderland, S.G., Taira, R.K., Brinkley, J.F. Requirements of a Web-Based Experiment Management System. AMIA Fall Symposium, 2000.

[2] Elliott, M. Considerations for Management of Laboratory Data. *2003 Scientific Computing & Instrumentation LIMS Guide, November 2003.*

[3] Zhu, J. M.S. thesis, Indiana University, Indianapolis, Indiana, June 2005.

[4] Kegelmeyer, P., Matarazzo C., Ambrosiano, J. Problem Solving Environment: Scientific Data Management. http://www.ca.sandia.gov/ASCI/pub/sc97/ (retrieved 1/13/2006), part of Sandia National Laboratories Accelerated Strategic Computing Initiative. http://www.ca.sandia.gov/

[5] McDowall, R.D., McDowall Consulting, Bromley, Kent, UK. Chromatography Data Systems III: Prospective Validation of a CDS. www.21cfrpart11.com/files/library/validation/cds_3.pdf (retrieved March 6, 2007)

[6] Title 21 Code of Federal Regulations (CFR) Part 820, and 61 Federal Register (FR) 52602.

[7] Risk-Based Validation of Commercial Off-the-Shelf Computer Systems. Published by Advanstar Communications in Journal of Validation Technology. May 2005, Vol. 11, No. 3.

[8] General Principles of Software Validation; Final Guidance for Industry Staff and FDA Staff. January 11, 2002.

[9] Tutorial - Computer System Validation. Lab Compliance. http://www.labcompliance.com/tutorial/csv/default.aspx?sm=d_b (retrieved February 25, 2008).

[10] Schumacher, W. Industry Guidelines for Computerized Systems Validation (GAMP, PDA Technical Reports). Training Course. Computerized System Validation in the Pharmaceutical Industry. Istanbul, 16-17 January 2003 (retrieved February 25, 2008).

[11] GAMP 4 Software Categories. Validation-Online.net. http://www.validation-online.net/GAMP-4.html (retrieved February 25, 2008).

[12] Search Software Quality. http://searchsoftwarequality.techtarget.com/sDefinition/0,,sid92_gci1243425,00.html

[13] Search Software Quality. http://searchsoftwarequality.techtarget.com/sDefinition/0,,sid92_gci1243430,00.html

[14] IEEE Standard Computer Dictionary: A Compilation of IEEE Standard Computer Glossaries; IEEE; New York, NY; 1990. Retrieved from "http://en.wikipedia.org/wiki/System_testing".

[15] CIO Mid-Market. http://searchcio-midmarket.techtarget.com/sDefinition/0,,sid183_gci836031,00.html

[16] Risk-Based Validation: What is it, and why do we do it? Life on the Pharm [online], January 17, 2005. http://life-on-the-pharm.blogspot.com/2005/01/risk-based-validation-what-is-it-and.html. (retrieved May 14, 2007).

[17] Huber, L. Qualification and validation of software and computer systems in laboratories. Part 4. Evaluation and validation of existing systems. *Accred Qual Assur* 1998 3:317-321.

[18] CODATA and the World Summit on the Information Society

[19] Rees, P. It's all down to interpretation. Scientific Computing World. February-March 2006 Issue [Online], http//:www.scientific-computing.com/scwfebmar06bridginggaps.html (retrieved April 5, 2006).

[20] Reed-Jones, N. Instrument Integration in Pharmaceutical QA Laboratories – Is it worth it? http://www.selectscience.net/jump/articles/3200 (retrieved 3/13/2006).

[21] Market profile: SDMS. AllBusiness [Online] http://www.allbusiness.com/periodicals/printArticle.asp?ID=562090 (retrieved 3/1/2006).

[22] 20th Feb, 2003, FDA Completes First Steps of its Broad Initiative to Improve Regulation of Pharmaceutical Manufacturing, www.fda.gov/bbs/topics/news/2003.

[23] Financial benefits of Linking Instruments to LIMS, technical ref sheet 029, CSols plc, http://www.csols.com (retrieved April 5, 2006).

[24] Agilent Technologies and Labtronics Partner to Provide Advanced Integration Solution for Cerity Enterprise Content Management (ECM). Labtronics [Online] February 2006, http://www.labtronics.com/news/Feb06_LLECM.htm (retrieved April 5, 2006).

[25] Laboratorytalk Editorial Team. Advanced integration for content management. Laboratorytalk [Online], February 24, 2006, http://www.laboratorytalk.com/news/laa/laa169.html (retrieved April 5, 2006).

[26] Laboratorytalk Editorial Team. Integrate RS232 instruments with SDMS. Laboratorytalk [Online], October 7, 2006, http://www.laboratorytalk.com/news/laa/laa127.html (retrieved April 5, 2006).

[27] Waters Corporation Web site, http://www.waters.com/WatersDivision/contentd.asp?watersit=EGOO-6KY2K4 (retrieved April 6, 2006).

[28] Robertson, S., Robertson, J. Mastering the Requirements Process; ACM Press, 1999

www.codata.org/wsis/CODATAandWSIS.pdf

[29] Waters Corporation. 2006. 01 February 2006 <http://www.waters.com/watersdivision/ContentD.asp?watersit=JDRS-5WJQ2W>

[30] Agilent Technologies. 2006. 16 February 2006 <http://www.chem.agilent.com/scripts/Pcol.asp?lpage=36989>

[31] Wikipedia. http://en.wikipedia.org.

[32] Amartus. 2006 07 March 2005 <http://www.amartus.com/sdms.html>

[33] Computer Compliance Inc. 2004 08 March 2006 <http://www.eflexion.com/home.asp>

[34] Abrevity. 2006 20 March 2005 <http://www.abrevity.com/filedata.html>

[35] Biotrue 2006 15 March 2006 <http://www.biotrue.net/site/main/>

[36] Toolkit project scope, requirements, and constraints for Agilent/SSI Assessment. October 24, 2004, Retrieved April 25, 2006.

APPENDICES

Appendix A - Volere Requirements Specification template

1.) The Purpose of the Product
    a. The user problem or background of the project
    b. Goals of the Product
2.) Client, Customer and Other Stakeholders
    a. The client is the person(s) paying for the development, and future owner of the delivered product
    b. The customer is the person who will buy the product
    c. Other stakeholders
3.) User of the Product
    a. The users of the product
    b. The priorities assigned to users
4.) Requirements Constraints
    a. Solution Constraints
    b. Implementation Environments
    c. Partner Applications
    d. Commercial off-the shelf software
    e. Anticipated workplace Environment
    f. How long do the developers have to build the product?
    g. What is the financial budget for the project?
5.) Naming Conventions and Definitions
6.) Relevant Facts
7.) Assumptions
8.) The Scope of the Product
    a. The context of the work
    b. Work Partitioning
    c. Product Boundary
9.) Functional and Data Requirements

a. Functional Requirements – Description of an action the product must take.
b. Data Requirements

10.) Look and Feel Requirements

11.) Usability Requirements
a. Ease of use
b. Ease of learning program

12.) Performance Requirements
a. Speed Requirements
b. Safety critical requirements
c. Precision Requirements
d. Reliability and availability requirements
e. Capacity Requirements

13.) Operational Requirements
a. Expected physical environment
b. Expected technological environment
c. Partner Applications

14.) Maintainability and Portability Requirements
a. How easy must it be to maintain this product?
b. Are there special conditions that apply to the maintenance of this product?
c. How portable must the program be?

15.) Security Requirements
a. Is the product confidential?
b. File integrity requirements
c. Audit Requirements

16.) Cultural and Political Requirements

17.) Legal Requirements
a. Does the product fall under the jurisdiction of any law?
b. Are there any standards with which we must comply?

18.) Open Issues

19.) Off-the-Shelf Solutions
    a. Is there a ready-made product that could be bought?
    b. Can ready-made components be used for this product?
    c. Is there something that we could copy?
20.) New Problems
    a. What problems could the new product cause in the current environment?
    b. Will the new development affect any of the installed systems?
    c. Will any of our existing users be adversely affected by the new development?
    d. What limitations exist in the anticipated implementation environment that may inhibit the new product?
    e. Will the new product create other problems?
21.) Tasks
    a. What steps have to be taken to deliver the product?
    b. Development phases
22.) Cutover
    a. What special requirements do we have to get the existing data, and procedures to work for the new product?
    b. What data has to be modified/translated for the new product?
23.) Risks
    a. What risks do you face when you develop this product?
    b. What contingency plans are you making?
24.) Costs – The more requirements the larger the cost
25.) User Documentation
    a. The plan for building the user documentation
26.) Waiting Room – What is next?

Appendix B - SDMS User Requirements

The following section lists the user requirements recorded for this thesis. Sections (from the Volere template) that are not listed are out of the scope of the project.

1 The Purpose of the Project

1a. The user problem or background of the project effort.

Today's pharmaceutical laboratories face a big challenge in determining how to handle the enormous amounts of data that are being generated. The rapid growth of scientific data today is forcing laboratories to move away from storing data in spreadsheets and small, non-robust databases toward a more advanced technology for data acquisition, storage, retrieval and collaboration among scientists. I want to define functional requirements for purchasing and installing a Scientific Data Management System (SDMS) in a pharmaceutical quality control laboratory to cope with the growth of data and collaboration issues among scientists.

A SDMS is an electronic "library" that collects, organizes, indexes, stores, archives and shares electronic records, from raw instrument data and reports to compliance records and others office documents (i.e. MS Office and others). SDMS also usually extract searchable metadata from each file and provides search capabilities and embedded viewers for many file types.

1b. Goals of the project.

The purpose of these requirements is to define the capabilities and characteristics to be used in designing or evaluating designs for a Scientific Data Management System (SDMS). This SDMS should meet the needs of the Pharmaceutical/Biotechnology industry (specifically research/development and quality control laboratories) for the next decade and provide improved performance relative to workflow management and efficiency.

Specific goals of this product are to improve business efficiency by:

- Improving scientist collaboration via electronically shared data
- Streamlining access to data sources

- Automating delivery of new information
- Reducing the time to get products to market

These requirements will provide direction and goals to be used by SDMS designers and manufacturers in developing a SDMS to meet the Pharmaceutical/Biotechnology industry needs.

2 Client, Customer and other Stakeholders

2a. Client

The client for the product is the global laboratory informatics group.

2b. Customer

The customers for the product are the Parenteral Quality Control laboratories.

2c. Other Stakeholders

a. Users (detailed in section 3)

b. Local IT support group

3 Users of the Product

3a. The hands-on users of the product

The scientists and technicians in the Quality Control laboratory organization are the main user group. The scientists and technicians are all experienced in using personal computers, and a wide variety of laboratory informatics applications including LIMS and CDS. Other users include the team supervisors, responsible scientists, quality control representatives and laboratory manager.

3b. The priorities assigned to users

User	Priority Rating	*Use Percentage
Scientist	Key User	90%
Technician	Key User	90%
QC Rep.	Secondary/Unimportant User	< 5%

RS	Secondary User	20%
Line Supervision	Key User	20%
Management	Secondary/Unimportant User	< 5%
Regional/Local Support	Key User	10%

* These percentages are estimated; actual usage percentage may vary.

3c. User participation

Users will assist in testing the requirements. Their customer satisfaction or dissatisfaction rating will be averaged and documented in the appropriate area of each requirement (where applicable).

3d. Maintenance users

Maintenance users for the product are the local IT department, the global informatics team, and the vendor.

4 Mandated Constraints

4a. Solution design constraints

a. User access to archived processed data will be determined by local policy.

b. User access to archived native instrument data will be determined by local policy.

c. A process will be established to process requests from external collaborators for archived data.

d. The retention times of archived data will be set in accordance with local and corporate retention policies.

e. For unsupported instrument data sources, users will provide file naming conventions and directory structures to facilitate creating SDMS templates.

f. An abbreviated verification process must be available to expedite change requests to systems serviced by SDMS.

g. A process must exist to allow authorized users to request un-scheduled file removal.

h. The file data capture and restore process must not interfere with the users' ability to acquire or process data.

i. The systems print data capture and view process must not interfere with the users' ability to acquire or process data.

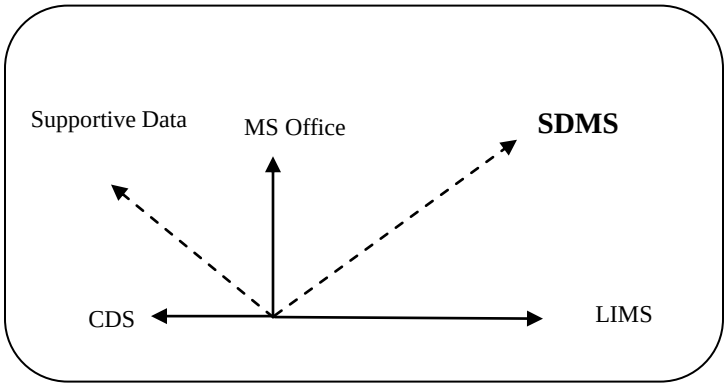
j. The systems print view process must not interfere with the users' ability to acquire or process data.

k. Network bandwidth will be sufficient to assure uninterrupted processing during execution of data archive, data restore, print data capture, or print view processes.

l. The system must use LDAP based authentication compatible with a simplified log-on process.

4b. Implementation environment of the current system

The SDMS system will be implemented within Quality Control laboratories (QCL). This document pertains only to the implementation of the SDMS that would be supported by QCL.



4c. Partner or collaborative applications

CDS

The CDS will collect and store chromatography test results, automatically linking instrument and sample processing methods with each result. It connects that data with the identity of the operator using the computer system. Each operator entry is automatically marked with a data and time stamp for each executed function. The SDMS

will use its automatic conversion capabilities to create JCAMP-DX public data standard files to view data from the CDS without the original application – for long-term data preservation

LIMS

The SDMS will store the result files that are generated in the process of completing testing on samples. LIMS will contain functionality to include a link to Print Data from test results and to retrieve and display these reports on demand. This functionality would be provided via a SDMS Web Service.

MS Office Suite/Supportive Data

SDMS will automatically capture data generated from any business application with a printer driver, including document, spreadsheet and presentation programs, and databases within minutes of its creation or change. This includes printed reports generated for scientist review—the software captures the actual content of the report, not just an image of that report.

4d. Off-the-shelf software

Not applicable.

4e. Anticipated workplace environment

The product will be run on the company intranet using a Citrix metaframe server. No other characteristics pose any known issues for product development or installation.

4f. How long do the developers have for the project?

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

4g. What is the financial budget for the project?

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

5 Naming Conventions and Definitions

Definitions and acronyms specific to this document are defined below.

Term/Acronym	Meaning
SDMS	Scientific Data Management System
ECM	Enterprise Content Management
CDS	Chromatography Data System
LIMS	Laboratory Information Management System
Native Instrument Data	File data, Raw data, and Files
Processed Data	Print data, Report data, and Results
API	Application Programming Interface
ID	Identification
LDAP	Lightweight Directory Access Protocol
Req.	Requirement
ER/ES	Electronic Records/Electronic Signatures
COA/COT	Certificate of Analysis/Certificate of Testing
RS	Responsible Scientist
QA	Quality Assurance
PPR&D	Pharmaceutical Product Research and Development
COTS	Commercial Off-the-Shelf
OJT	On-the-job training
ROI	Return on Investment
LAN	Local Area Network

6 Relevant Facts and Assumptions

6a. Factors that have an effect on the product, but are not mandated requirements constraints.

This section is out of the scope of this set of requirements, because the requirements are being created “post” product purchase in order to show requirements gathering for a SDMS.

6b. Assumptions that the team is making about the project

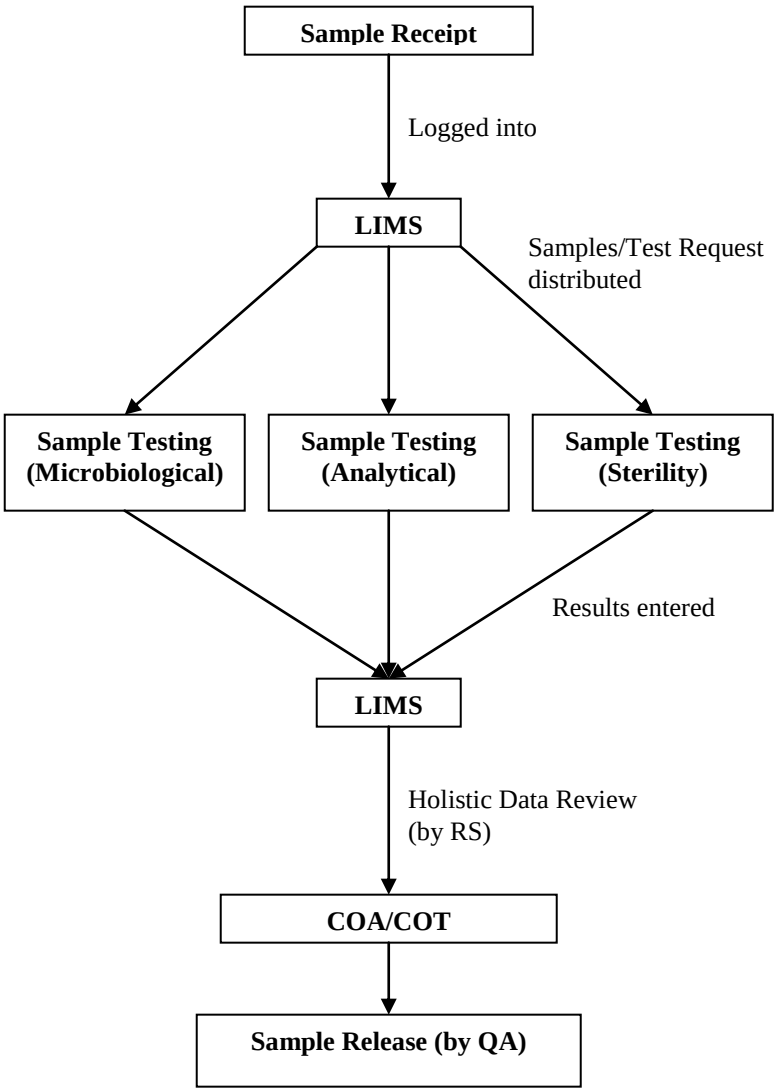
This section is out of the scope of this set of requirements, because the requirements are being created “post” product purchase in order to show requirements gathering for a SDMS.

7 The Scope of the Work

7a. The current situation

Currently, there is no integrated system with the functionality of a SDMS. We have several systems that act independently of each other. These include LIMS, CDS, Office Applications, and other laboratory instruments and reporting tools.

7b. The context of the work.



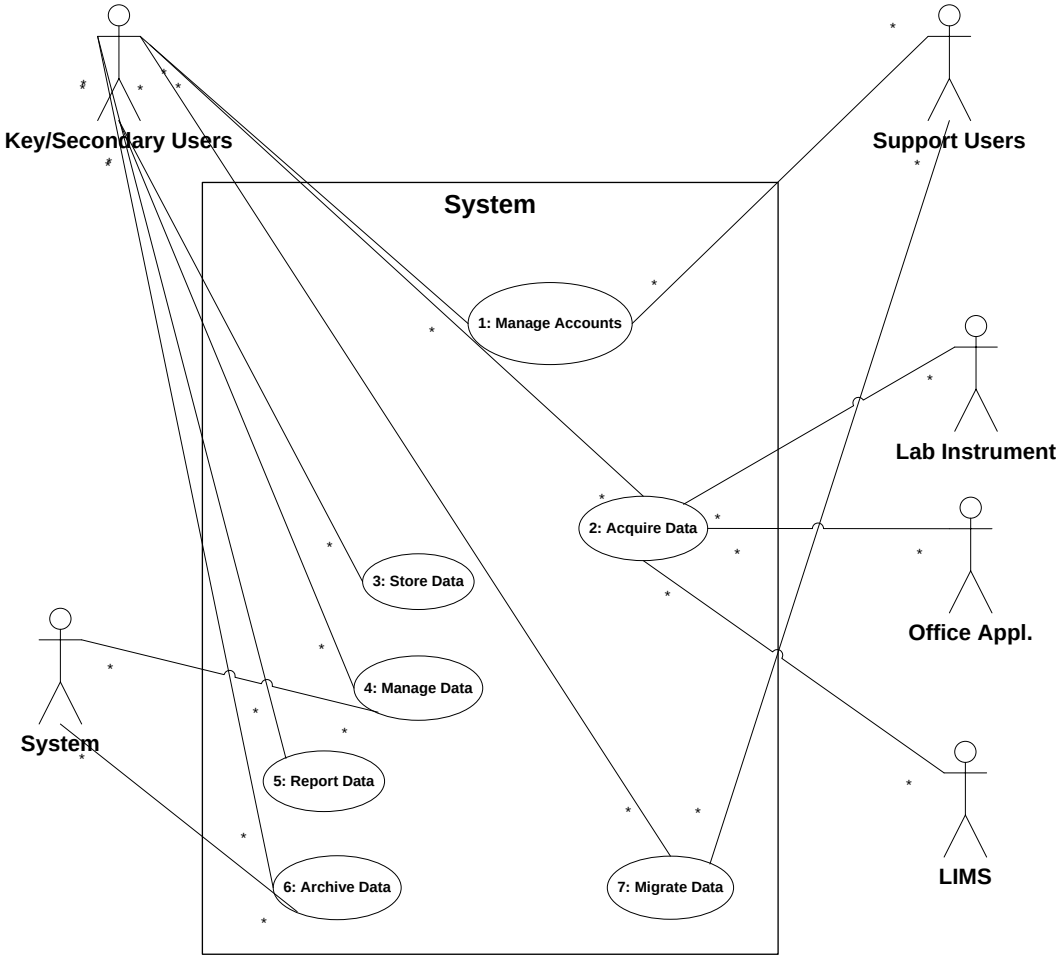
7c. Work partitioning

Business Event List

Event Name	**Input & Output**
Samples and Test Request submitted	Sample Test Request (in)
Lab personnel logs samples into LIMS	Samples logged in (in)
Sample Test Requests distributed to lab personnel.	Test Requests distributed (out)
Analysts test samples.	Results generated from disparate equipment (out)
Analysts record results.	Results entered in LIMS (in)
RS's perform holistic review of data.	COA/COT prepared (out)
QA releases samples	Samples shipped (out)

8 The Scope of the Product

8a Product Boundary



8b Product use case list

Use Case Information

Use Case ID:	01
Use Case Name:	Manage Accounts
Use Case Description:	Use case describes the functionality for user and system-level processes related to account management.
Pre-Conditions:	User has corporate LAN account

Scenario Information

Scenario:	A user logs into the system
Scenario Number	Sc01
Description:	This scenario proves that a user is able to access the system within the defined business rules.
Successful Outcome:	User successfully logs into the system.
Failed Outcome:	The system does not allow user to log in.
Primary Actor(s):	Scientists, Technicians, Line Supervisors, Regional or Local Support
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Scenario:	A support user create or modifies a user account
Scenario Number	Sc02
Description:	This scenario proves that a support user is able to create or modify a user account on the system within defined business rules.
Successful Outcome:	A support user successfully creates and modifies an end user's account.
Failed Outcome:	A support user is unable to create and/or modify an end user's account.
Primary Actor(s):	Regional Support, Local Support
Secondary Actor(s):	Not applicable

Scenario:	A user needs to reset their password
Scenario Number	Sc03
Description:	This scenario proves that a user is able to reset their password.
Successful Outcome:	The system allows a user to reset their account.
Failed Outcome:	The system does not allow a user to reset their account.
Primary Actor(s):	All Users
Secondary Actor(s):	Not applicable

Scenario:	A support user needs to reset a user's password
Scenario Number	Sc04
Description:	This scenario proves that a support user is able to reset a user's password.
Successful Outcome:	The system allows a support user to reset an end user's account.
Failed Outcome:	The system does not allow a support user to reset an end user's account.
Primary Actor(s):	Regional Support, Local Support
Secondary Actor(s):	Not applicable

Use Case Information

Use Case ID:	UC02
Use Case Name:	Acquire Data
Use Case Description:	Use case describes the functionality for user and system-level processes related to data acquisition.
Pre-Conditions:	Instruments are connected

Scenario Information

Scenario:	The system acquires instrument data from a laboratory instrument.
Scenario Number	Sc05
Description:	This scenario proves that the system permits acquisition of raw data from laboratory instruments within defined business rules.
Successful Outcome:	The system acquires data from an external instrument in its native format.
Failed Outcome:	The system is unable to acquire external instrument data or modifies the original file format after being acquired.
Primary Actor(s):	Laboratory Instrument
Secondary Actor(s):	Not applicable

Scenario:	The system acquires data from an office application.
Scenario Number	Sc06
Description:	This scenario proves that the system permits acquisition of native instrument or file data from an office application within defined business rules.
Successful Outcome:	The system acquires data from an external office application in its native format.
Failed Outcome:	The system is unable to acquire external office application data or modifies the original file format after being acquired.
Primary Actor(s):	Specific Office Application (i.e. MS Word)
Secondary Actor(s):	Not applicable

Scenario:	The system acquires data from a LIMS.
Scenario Number	Sc07
Description:	This scenario proves that the system permits acquisition of native instrument or file data from a LIMS application within defined business rules.
Successful Outcome:	The system acquires data from a LIMS in its native format.
Failed Outcome:	The system is unable to acquire LIMS data or modifies the original data format after being acquired.
Primary Actor(s):	LIMS
Secondary Actor(s):	Not applicable

Use Case Information

Use Case ID:	UC03
Use Case Name:	Store Data
Use Case Description:	Use case describes the functionality for user and system-level processes related to data storage.
Pre-Conditions:	User is logged in; the data is collected.

Scenario Information

Scenario:	A user stores processed data to a secure location.
Scenario Number	Sc08
Description:	This scenario proves that the system permits secured storage of captured data within defined business rules.
Successful Outcome:	Captured data is storable on a secured server.
Failed Outcome:	Captured data is not storable on a secured server.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Use Case Information

Use Case ID:	UC04
Use Case Name:	Manage Data
Use Case Description:	Use case describes the functionality for user and system-level processes related to managing captured data.
Pre-Conditions:	User is logged in; Data is already captured and stored.

Scenario Information

Scenario:	The system indexes native instrument data using metadata tags.
Scenario Number	Sc09
Description:	This scenario proves that the system permits indexing of captured data using metadata tags.
Successful Outcome:	The system indexes captured data based on available metatags, defined by the user.
Failed Outcome:	The system does not index captured data based on available metatags, defined by the user.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Scenario:	The system allows captured data to be converted to a data exchange format.
Scenario Number	Sc10
Description:	This scenario proves that the system permits the conversion of captured data to a data exchange viewer format within defined business rules.
Successful Outcome:	A user converts a data file to a data exchange viewer format, such as JCAMP-DX.
Failed Outcome:	The conversion is a data file to a human-viewable data exchange format is unsuccessful.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Scenario:	The system preserves the original format of captured data.
Scenario Number	Sc11
Description:	This scenario proves that the system preserves the original format of captured data.
Successful Outcome:	A user cannot tell the difference between the data from its original instrument, or application and the file captured in the SDMS.
Failed Outcome:	A user sees a difference between a captured data file in the SDMS and the original file format.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Scenario:	A user opens and views instrument data without restoring them.
Scenario Number	Sc12
Description:	This scenario proves that the system permits the viewing of captured data without having to restore the file or application source.
Successful Outcome:	A user successfully opens an instrument data file, stored within the SDMS, without restoring the externally connected application.
Failed Outcome:	The instrument data file or application has to be restored to view a selected file.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Use Case Information

Use Case ID:	UC05
Use Case Name:	Report Data
Use Case Description:	Use case describes the functionality for reporting data, whether to a screen or to a printer.
Pre-Conditions:	User is logged in; Data is collected

Scenario Information

Scenario:	A user formats a report
Scenario Number	Sc13
Description:	This scenario proves that a user is able to format a report within defined business rules.
Successful Outcome:	A user formats a report summary, presentation, electronic submission, or publication.
Failed Outcome:	A user cannot format a report summary, presentation, electronic submission, or publication.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Scenario:	A user creates a report
Scenario Number	Sc14
Description:	This scenario proves that a user is able to create a report within defined business rules.
Successful Outcome:	A user creates a report summary, presentation, electronic submission, or publication.
Failed Outcome:	A user cannot create a report summary, presentation, electronic submission, or publication.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Scenario:	A user prints a report
Scenario Number	Sc15
Description:	This scenario proves that a user is able to print a report within defined business rules.
Successful Outcome:	A user prints a report summary, presentation, electronic submission, or publication.
Failed Outcome:	A user cannot print a report summary, presentation, electronic submission, or publication.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers

Use Case Information

Use Case ID:	UC06
Use Case Name:	Archive Data
Use Case Description:	Use case describes the functionality for user and system-level processes related to data archival.
Pre-Conditions:	User is logged in; data is captured and metadata tags are defined

Scenario Information

Scenario:	The system allows for scheduled archiving of data
Scenario Number	Sc16
Description:	This scenario proves that the system permits the automatic archival of captured data based on a user-defined schedule.
Successful Outcome:	The system allows a configurable scheduled archive, without human intervention.
Failed Outcome:	The system does not allow for a configurable scheduled archive, without human intervention.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers, Regional Support, Local Support

Scenario:	A system indexes archived, processed data.
Scenario Number	Sc17
Description:	This scenario proves that the system will index archived, processed data based on defined metadata tags.
Successful Outcome:	The system indexes archived data based on available metatags, defined by the user.
Failed Outcome:	The system does not index archived data based on available metatags, defined by the user, or the indexing is incorrect.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers, Regional Support, Local Support

Use Case Information

Use Case ID:	UC07
Use Case Name:	Migrate Data
Use Case Description:	Use case describes the functionality for user and system-level processes related to data migration.
Pre-Conditions:	User is logged in; Data is collected.

Scenario Information

Scenario:	Data from a previous version the software needs migrating to current version.
Scenario Number	Sc18
Description:	This scenario proves that the system permits the migration of captured data between different software versions.
Successful Outcome:	A Support user migrates data from a previous software version to the current version, and the data is not altered.
Failed Outcome:	A Support user cannot migrate data from a previous software version to the current version, and the data is not altered.
Primary Actor(s):	Regional Support, Local Support
Secondary Actor(s):	Not applicable

Scenario:	Data from a previous version the software needs to be viewed.
Scenario Number	Sc19
Description:	This scenario proves that the system permits the viewing of migrated data in its original format, and will all the original content.
Successful Outcome:	Any authorized user can view data from a previous software version, and the data matches the original.
Failed Outcome:	An authorized user cannot view data from a previous software version, or the data does not match the original.
Primary Actor(s):	Scientists, Technicians, Line Supervisors
Secondary Actor(s):	QC Representatives, Responsible Scientists, Managers, Regional Support, Local Support

9 Functional and Data Requirements

9a. Functional Requirements.

Platform

Requirement #:	1	Requirement Type:	9	Event/use case #:	N/A
Description:	The system must support instruments based on Unix.				
Rationale:	Many instruments have a Unix platform.				
Source:	PPR&D				
Fit Criterion:	The system successfully interfaces with instruments based a UNIX configuration.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	Current department topology				
History:	Created February 27, 2006				

Requirement #:	2	Requirement Type:	9	Event/use case #:	N/A
Description:	The system must support data capture from instruments based on Windows NT SP6a or newer.				
Rationale:	Windows NT SP6a is the oldest software version supported.				
Source:	PPR&D				
Fit Criterion:	The system successfully captures data from instruments based on Windows NT SP6a or newer.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	Current department topology				
History:	Created February 27, 2006				

Requirement #:	3	Requirement Type:	9	Event/use case #:	N/A
Description:	The system must support previewing and restore of data on instruments based on Windows 2000 SP4 or newer.				
Rationale:	Windows 2000 SP4 is the oldest version supported for this operation.				
Source:	PPR&D				
Fit Criterion:	Data can be previewed and restored from instruments based on Windows 2000 SP4 or newer.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	Current department topology				
History:	Created February 27, 2006				

Requirement #:	4	Requirement Type:	9	Event/use case #:	N/A
Description:	The system must support instruments based on MacOS X (10.2 or newer).				
Rationale:	Some instruments have a MacOS software.				
Source:	PPR&D				
Fit Criterion:	The system successfully supports instruments based on MacOS X (10.2 or newer).				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	Current department topology				
History:	Created February 27, 2006				

Requirement #:	5	Requirement Type:	9	Event/use case #:	N/A
Description:	The system must support print capture from a Citrix based environment.				
Rationale:	Many of the applications are run via a Citrix metaframe.				
Source:	PPR&D				
Fit Criterion:	Data can be printed to the SDMS from an application in a Citrix environment.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	Current department topology				
History:	Created February 27, 2006				

Native Instrument Data

Requirement #:	6	Requirement Type:	9	Event/use case #:	UC02
Description:	The system must pull native instrument data from a targeted file share and save it to a secure server.				
Rationale:	Data should be maintained in a secure environment at all times.				
Source:	PPR&D				
Fit Criterion:	Native instrument data can be pulled from a file share and saved to the server.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	Refer to the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	7	Requirement Type:	9	Event/use case #:	UC04
Description:	The SDMS system must be able to restore native instrument file data.				
Rationale:	The original file should be kept in tact. File modifications should be captured to see differences in versions.				
Source:	PPR&D				
Fit Criterion:	A new file data version will be added to the SDMS with the updated information and showing the modified date and time.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	8	Requirement Type:	9	Event/use case #:	UC04
Description:	The system will preserve the original, native instrument data document formats.				
Rationale:	The original file should be kept in tact. File modifications should be captured to see differences in versions.				
Source:	PPR&D				
Fit Criterion:	A new file data version will be added to the SDMS with the updated information and showing the modified date and time.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	9	Requirement Type:	9	Event/use case #:	UC04
Description:	The system will allow captured native instrument files to be converted into JCAMP-DX.				
Rationale:	JCAMP-DX is a public conversion standard used for most spectroscopy data; LC/MS/MS standard in progress.				
Source:	PPR&D				
Fit Criterion:	A chosen file is successfully converted into JCAMP DX format.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	10	Requirement Type:	9	Event/use case #:	UC04
Description:	The system will allow the user to open and view specific captured native instrument files without restoring them.				
Rationale:	Minimizing time and the need for manually launching the original application.				
Source:	PPR&D				
Fit Criterion:	User successfully opens a previously captured file without having to restore it.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	11	Requirement Type:	9	Event/use case #:	UC06
Description:	The system must allow the specification of a retention period for archived native instrument data.				
Rationale:	Per regulations and SOPs, data is required to be retained for a specified period.				
Source:	PPR&D				
Fit Criterion:	A data archival retention period specification can be set for any instrument data.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created March 7, 2006				

Archive Data

Requirement #:	12	Requirement Type:	9	Event/use case #:	UC06
Description:	The system must allow for a configurable scheduled archive, without human intervention of native instrument data from designated folders.				
Rationale:	Data must be able to be archived on a schedule so that users don't have to remember to do so.				
Source:	PPR&DPRR&D				
Fit Criterion:	The user is allowed to schedule an automatic archive.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	13	Requirement Type:	9	Event/use case #:	UC06
Description:	The system must allow for archived native instrument data to be restored with its original directory structure.				
Rationale:	Archived data must be available in the original structure for audits.				
Source:	PPR&D				
Fit Criterion:	A user can successfully restore archived data back to the original format.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	14	Requirement Type:	9	Event/use case #:	UC06
Description:	The system must allow for removal of successfully archived native instrument data from its original location on a configurable schedule.				
Rationale:	Data must be able to be archived on a schedule so that users don't have to remember to do so.				
Source:	PPR&D				
Fit Criterion:	An instrument data file is automatically archived based on a set schedule.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	15	Requirement Type:	9	Event/use case #:	UC06
Description:	The system must allow for native instrument data archive schedules to be instrument specific.				
Rationale:	To be able to coordinate archival by instrument or instrument type.				
Source:	Lilly PPR&D				
Fit Criterion:	An archive schedule is successfully set for a data file by instrument or instrument type.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	16	Requirement Type:	9	Event/use case #:	UC06
Description:	The system will assure that native instrument data locked for update or acquisition during a scheduled archive will be captured during the next scheduled archive.				
Rationale:	To avoid missing any data requiring archival.				
Source:	PPR&D				
Fit Criterion:	Locked data file will not be archived during scheduled archive. Unlocked data file will be archived during scheduled archive.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	17	Requirement Type:	9	Event/use case #:	UC06
Description:	The system must allow for successfully archived native instrument data to be moved to an off-line secure storage location.				
Rationale:	Prevention of the loss of raw data.				
Source:	PPR&D				
Fit Criterion:	A data file is successfully stored offline and matches the file when online.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	18	Requirement Type:	9	Event/use case #:	UC06
Description:	The archived SDMS native instrument data must be searchable				
Rationale:	May have to be retrieved for further analysis or for use in an audit.				
Source:	PPR&D				
Fit Criterion:	An instrument data file is retrieved in list of results when search for using metadata.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	19	Requirement Type:	9	Event/use case #:	UC06
Description:	The system will perform archive and restore functions concurrent with data acquisition.				
Rationale:	To save time and allow for multitasking.				
Source:	PPR&D				
Fit Criterion:	A user successfully archives and/or restores a data file while also acquiring new data simultaneously.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Metadata

Requirement #:	20	Requirement Type:	9	Event/use case #:	UC04
Description:	The SDMS system will use metadata tags to describe the native instrument data.				
Rationale:	Makes searching easier and reduces storage space.				
Source:	PPR&D				
Fit Criterion:	A user can query for and retrieve instrument data via its metadata tag.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	21	Requirement Type:	9	Event/use case #:	UC04
Description:	The archived SDMS native instrument data must be indexed using available metatags.				
Rationale:	Makes searching easier and reduces storage space.				
Source:	PPR&D				
Fit Criterion:	A user can query for and retrieve instrument data via its metadata tag.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	22	Requirement Type:	9	Event/use case #:	UC04
Description:	The system must display instrument metadata in a human readable format.				
Rationale:	The data is readable by the human eye.				
Source:	PPR&D				
Fit Criterion:	The metadata is visible and readable by a user.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Processed Data

Requirement #:	23	Requirement Type:	9	Event/use case #:	UC03
Description:	The system must allow processed data to be submitted to a secure, accessible location.				
Rationale:	Data should remain secure at all times to avoid unauthorized manipulation.				
Source:	PPR&D				
Fit Criterion:	Only an authorized user can access the raw data file from the server.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	24	Requirement Type:	9	Event/use case #:	UC06
Description:	The archived SDMS processed data must be indexed.				
Rationale:	Provides easier accessibility and searching.				
Source:	PPR&D				
Fit Criterion:	The processed data has a unique index (i.e. metadata tag).				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	25	Requirement Type:	9	Event/use case #:	UC04
Description:	The archived SDMS processed data must be searchable.				
Rationale:	Easier retrieval of data.				
Source:	PPR&D				
Fit Criterion:	User retrieves a processed data file by searching the index.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	26	Requirement Type:	9	Event/use case #:	UC04
Description:	The system must display processed data in a human readable format.				
Rationale:	The data is readable by the human eye.				
Source:	PPR&D				
Fit Criterion:	The processed data is visible and readable by a user.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	27	Requirement Type:	9	Event/use case #:	UC04
Description:	The system will preserve the original appearance of the captured process data.				
Rationale:	To avoid the assumption (by a regulatory body) of altered data files.				
Source:	PPR&D				
Fit Criterion:	Appearance and file type of the data looks the same as it did originally when viewed by a user.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	28	Requirement Type:	9	Event/use case #:	UC02
Description:	The system will perform print data capture and display functions concurrent with data acquisition.				
Rationale:	Reduce time to complete tasks.				
Source:	PPR&D				
Fit Criterion:	Both the instrument data acquisition process and the SDMS print data submission process will run concurrently, and successfully, to completion.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	29	Requirement Type:	9	Event/use case #:	UC04
Description:	The system must allow captured processed information to be reused by other applications.				
Rationale:	Prevents reprocessing data, which reduces rework and time.				
Source:	PPR&D				
Fit Criterion:	A user can extract already processed data and use it for their needs without having to reprocess the data.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	30	Requirement Type:	9	Event/use case #:	UC04
Description:	The system will allow a user to search for similar records based upon a set of metadata tag values.				
Rationale:	Prevents having to perform multiple searches for similar information.				
Source:	PPR&D				
Fit Criterion:	A user successfully performs a search for similar processed data, group by its metadata tags.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	31	Requirement Type:	9	Event/use case #:	UC04
Description:	The system must provide a method to extract data from captured processed data and present it in a human readable format.				
Rationale:	To review previously processed data at any time during its record retention.				
Source:	PPR&D				
Fit Criterion:	A user successfully extracts previously processed data and presents it in a format readable to the human eye.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Requirement #:	32	Requirement Type:	9	Event/use case #:	UC03
Description:	The system must allow the specification of a retention period for captured processed data.				
Rationale:	Per regulations and SOPs, data is required to be retained for a specified period.				
Source:	PPR&DPR&D				
Fit Criterion:	The retention period for a processed data file is set successfully.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Reference the defined Use Case.				
History:	Created March 7, 2006				

Migrated Data

Requirement #:	33	Requirement Type:	9	Event/use case #:	UC07
Description:	The system must allow data from earlier versions to be migrated to the SDMS current version.				
Rationale:	Software is constantly being improved; users need a way to view data from an older version in the latest release.				
Source:	PPR&D				
Fit Criterion:	Data is successfully backed up and restored in the latest version without alteration.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created February 27, 2006				

Requirement #:	34	Requirement Type:	9	Event/use case #:	UC07
Description:	The system must allow the viewing of print data from earlier versions of the SDMS.				
Rationale:	Software is constantly being improved; users need a way to view data from an older version in the latest release.				
Source:	PPR&D				
Fit Criterion:	User successfully views print data from an earlier version of the software.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Refer to the defined Use Case.				
History:	Created February 27, 2006				

9b. Data requirements.

This section is adequately dealt with by defining the terms in the described in section 5, and by the use case diagram and work context in section 7.

10 Look and Feel Requirements

10a. The interface

Requirement #:	35	Requirement Type:	10	Event/use case #:
Description:	The system must have an API or similar toolkit to enable integration with other applications.			
Rationale:	Integrating the SDMS with other applications and instruments is the main reason for having the system.			
Source:	PPR&D			
Fit Criterion:	The SDMS Software Developers Kit CD will be available, installs successfully, and passes the current SDMS Software Development Kit installation qualification with no failures. The optional Software Developers Kit will connect to the specified server and display the specified information.			
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD	
Dependencies:	N/A	Conflicts:	N/A	
Supporting Materials:	N/A			
History:	Created February 23, 2006			

Requirement #:	36	Requirement Type:	10	Event/use case #:	
Description:	The system must provide an interface for external clients to search for and view processed data reports in the SDMS based on metadata tags.				
Rationale:	External partners may need access to view data reports. Easier than mailing hardcopies or faxes.				
Source:	PPR&D				
Fit Criterion:	A report is successfully generated by searching for information based on metadata tags.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created February 23, 2006				

10b. The style of the product

Requirement #:	37	Requirement Type:	10	Event/use case #:	N/A
Description:	The product should have an appearance suitable for the priority users specified in section 3b to understand.				
Rationale:	Different levels of experience and understanding for each user.				
Source:	J. Heyward				
Fit Criterion:	Any user listed in section 3b can navigate the software with minimal error or confusion in within 2 hours of training.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 6, 2006				

11 Usability and Humanity Requirements

11a. Ease of use

Requirement #:	38	Requirement Type:	11	Event/use case #:	N/A
Description:	The product shall easy to use by a novice or advanced user.				
Rationale:	Not all users have advanced computer software skills.				
Source:	J. Heyward				
Fit Criterion:	One month's use of the product shall result in a total error rate of less than approximately 5%.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 1, 2006				

11b. Personalization and internationalization requirements

Requirement #:	39	Requirement Type:	11	Event/use case #:	N/A
Description:	The product shall be customizable for a chosen language.				
Rationale:	Lilly has sites in different countries and the product may eventually being deployed globally.				
Source:	J. Heyward				
Fit Criterion:	International users can convert to their native language.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 1, 2006				

Requirement #:	40	Requirement Type:	11	Event/use case #:	N/A
Description:	The product shall allow users to save personal preferences.				
Rationale:	Users have the opportunity to participate more closely with the organization, as well as have their own personal user experience.				
Source:	J. Heyward				
Fit Criterion:					
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 6, 2006				

11c. Ease of learning.

Requirement #:	41	Requirement Type:	11	Event/use case #:	N/A
Description:	A trained user shall be able to be productive within a short time.				
Rationale:	Turnaround time for product release, etc. in a quality control lab is crucial.				
Source:	J. Heyward				
Fit Criterion:	75% of a test panel shall successfully complete specified assigned tasks within 2 hours after completing 8 hours of training.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 1, 2006				

11d. Understandability and Politeness requirements.

Requirement #:	42	Requirement Type:	11	Event/use case #:	N/A
Description:	The product shall use symbols and words that are naturally understandable in the user's domain.				
Rationale:	Users should not have to learn terms that are intrinsically unique to the product's internal construction or from another field.				
Source:	J. Heyward				
Fit Criterion:	75% of a test panel understands the symbols and words used in the software upon their first use.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 1, 2006				

11e. Accessibility requirements.

Not applicable

12 Performance Requirements

12a. Speed and latency requirements

Although this section may be important in the future, it has not been determined what speed and latency requirements are needed, if any.

12b. Safety critical requirements

Not applicable

12c. Precision or accuracy requirements

Not applicable

12d. Reliability and Availability requirements

Requirement #:	43	Requirement Type:	12	Event/use case #:	N/A
Description:	The SDMS should be available for use 24 hours per day, 365 days per year.				
Rationale:	Some labs operate 24/7.				
Source:	J. Heyward				
Fit Criterion:	Cannot be tested; system design has the option to meet the requirement demands.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 27, 2006				

Requirement #:	44	Requirement Type:	12	Event/use case #:	N/A
Description:	The SDMS should achieve 99% up time.				
Rationale:	Users will not have access to data stored in SDMS if it is down.				
Source:	J. Heyward				
Fit Criterion:	Less than or equal to 1% downtime on average per year.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	March 27, 2006				

12e. Robustness or fault tolerance requirements

Requirement #:	45	Requirement Type:	12	Event/use case #:	N/A
Description:	The system must be able to operate in a local mode when its server connection is lost.				
Rationale:	To ensure all services are available during abnormal occurrences.				
Source:	J. Heyward				
Fit Criterion:	Cannot be tested; system design has the option to meet the requirement demands.				
Customer Satisfaction:	N/A	Customer Dissatisfaction:	N/A		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created April 10, 2006				

12f. Capacity requirements

Requirement #:	46	Requirement Type:	12	Event/use case #:	N/A
Description:	The product shall be available for approximately 150 simultaneous users between 6:00am – 5:00pm. Maximum loading periods between 5:00pm and 6:00am will be approximately 10-15.				
Rationale:	To ensure the product is capable of processing the expected volumes of data.				
Source:	J. Heyward				
Fit Criterion:	Cannot be tested as part of this project; system should be built to meet the requirement demands.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created April 10, 2006				

12g. Scalability or extensibility requirements

Although this is an important section to consider, there are currently no known plans to increase production or staff.

12h. Longevity requirements

Requirement #:	47	Requirement Type:	12	Event/use case #:	N/A
Description:	The product is expected to operate within the maintenance budget for a minimum of 5 years.				
Rationale:	Minimize cost to build or purchase a new product and to maximize the ROI.				
Source:	J. Heyward				
Fit Criterion:	Cannot be tested; system design has the option to meet the requirement demands.				
Customer Satisfaction:	N/A	Customer Dissatisfaction:	N/A		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	Department budget				
History:	Created March 27, 2006				

13 Operational Requirements

13a. Expected physical environment

Requirement #:	48	Requirement Type:	13	Event/use case #:	N/A
Description:	The product shall be installed in a data center and deployed via a Citrix metaframe server.				
Rationale:	See section 4e.				
Source:	J. Heyward				
Fit Criterion:	The system in				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 27, 2006				

13b. Expected technological environment

Not known at this time.

13c. Partner applications

Refer to section 4b.

13d. Productization requirements

Requirement #:	49	Requirement Type:	13	Event/use case #:	N/A
Description:	The product shall be installed by the vendor.				
Rationale:	To ensure the vendor warranty is not compromised.				
Source:	J. Heyward				
Fit Criterion:	Vendor completes IQ/OQ to customer satisfaction.				
Customer Satisfaction:	N/A	Customer Dissatisfaction:	N/A		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	J. Heyward				

History:	Created March 27, 2006

14 Maintainability and Support Requirements

14a. Maintenance requirements

Requirement #:	50	Requirement Type:	14	Event/use case #:	N/A
Description:	The product should be able to be maintained by local IT users that are not the original developers of the product.				
Rationale:	Once the vendor installs the software and trains users, it will be the "customer's" responsibility to maintain the product.				
Source:	J. Heyward				
Fit Criterion:	Local IT personnel have security privileges to access vendor code and any other functions needed to maintain the software.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 28, 2006				

14b. Special conditions that apply to the maintenance of the product

SDMS maintenance activities will be governed by local release management procedures.

14c. Supportability requirements

SDMS support activities will be governed by local release management procedures.

14d. Adaptability requirements

Requirement #:	51	Requirement Type:	14	Event/use case #:	N/A
Description:	The product shall be translated into various foreign languages.				
Rationale:	Lilly has sites in different countries and the product may eventually being deployed globally.				
Source:	J. Heyward				
Fit Criterion:	International users can convert to their native language.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:	TBD		
Dependencies:	N/A	Conflicts:	N/A		
Supporting Materials:	N/A				
History:	Created March 28, 2006				

14e. Installation requirements

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS. However, the SDMS installation process will be governed by local installation procedures.

15 Security Requirements

15a. Access requirements

Requirement #:	52	Requirement Type:	15	Event/use case #:	UC01
Description:	System administrators must be able to create, modify, disable, and deactivate user accounts.				
Rationale:	Federal Regulation				
Source:	J. Heyward				
Fit Criterion:	Administrators can create, modify, disable, and deactivate user accounts.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created March 6, 2006				

Requirement #:	53	Requirement Type:	15	Event/use case #:	UC01
Description:	The system must be limited to authorized individuals.				
Rationale:	Federal Regulation				
Source:	21 CFR Part 11.300d, 11.10d				
Fit Criterion:	System allows access to user with valid ID and related password. System denies user access with invalid system ID and a valid user password.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created March 6, 2006				

Requirement #:	54	Requirement Type:	15	Event/use case #:	UC01
Description:	The system must lock a user out after three consecutive login attempts.				
Rationale:	Security threat				
Source:	21 CFR Part 11.300b				
Fit Criterion:	System disables user ID after the third login attempt. Message box is displayed stating the user's account has been locked out.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created March 6, 2006				

Requirement #:	55	Requirement Type:	15	Event/use case #:	UC01
Description:	User passwords must be changed every 60 days.				
Rationale:	Passwords should be changed periodically for security reasons.				
Source:	21 CFR Part 11.300b				
Fit Criterion:	User is prompt to change password after 60 days.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 13, 2006				

15b. Integrity requirements

Requirement #:	56	Requirement Type:	15	Event/use case #:	UC01
Description:	Users must be able to change their own passwords.				
Rationale:	Passwords are private, thus an unauthorized user should not be able to change another user's password.				
Source:	21 CFR Part 11.300a				
Fit Criterion:	User successfully changes his/her own password. The system does not allow the user to change another person's password.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 13, 2006				

Requirement #:	57	Requirement Type:	15	Event/use case #:	N/A
Description:	When changing passwords, users will be prevented from re-using their current password.				
Rationale:	This minimizes the chances of an unauthorized user finding out another user's password.				
Source:	21 CFR Part 11.300b				
Fit Criterion:	User cannot use a previous password. An error is generated. User is prompted to enter a valid/unique password.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 13, 2006				

15c. Privacy requirements

Requirement #:	58	Requirement Type:	15	Event/use case #:	N/A
Description:	Passwords must not be displayed or printed when entered.				
Rationale:	Possible security breach.				
Source:	21 CFR Part 11.300d				
Fit Criterion:	A users' password is encrypted when entered.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 13, 2006				

Requirement #:	59	Requirement Type:	15	Event/use case #:	N/A
Description:	Passwords must be maintained in a secure manner.				
Rationale:	Possible security breach.				
Source:	21 CFR Part 11.300d				
Fit Criterion:	Users' password cannot be seen by an administrative. User has no access to system ID information of other system users.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 13, 2006				

15d. Audit requirements

Requirement #:	60	Requirement Type:	15	Event/use case #:	N/A
Description:	The system must be capable of displaying an access roster for periodic review by an administrator or local IT.				
Rationale:	Administrator's should periodically review the current roster to ensure that: Authorized users are all current personnel Authorized users have the correct access level Authorized users have completed all initial and subsequent training needed to complete their assigned tasks.				
Source:	21 CFR Part 11.10i				
Fit Criterion:	System displays an access roster (upon request) showing all system users and their assigned privileges.				
Customer Satisfaction:	TDB	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 13, 2006				

15e. Immunity requirements

Requirement #:	61	Requirement Type:	15	Event/use case #:	N/A
Description:	The system must be able to protect itself from malicious interference (i.e. viruses, worms, etc.).				
Rationale:	Risk to corrupting all data stored in the SDMS.				
Source:	J. Heyward				
Fit Criterion:	Antivirus software is installed on the server running the SDMS or on the client PC running the software (if applicable).				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR 21 Part 11.10c				
History:	Created March 13, 2006				

16 Cultural and Political Requirements

16a. Cultural requirements

Not applicable.

16b. Political requirements

Not applicable.

17 Legal Requirements

17a. Compliance requirements

Requirement #:	62	Requirement Type:	17	Event/use case #:	N/A
Description:	The system must be able to discern invalid or altered records.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10a				
Fit Criterion:	System prompts user to save the changes before it will close. The system will not allow the invalid data to be added to the template.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	63	Requirement Type:	17	Event/use case #:	N/A
Description:	The system must be able to generate accurate and complete copies of records in both human-readable and electronic form for review and copying.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10b				
Fit Criterion:	Record is recalled in human readable form and is identical to the hardcopy record.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	64	Requirement Type:	17	Event/use case #:	N/A
Description:	The records must be protected to ensure that they are readily retrievable throughout the applicable retention period.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10c				
Fit Criterion:	A data file stored in offline storage can be retrieved at any time during the record retention period.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	65	Requirement Type:	17	Event/use case #:	N/A
Description:	The system must provide secure computer-generated, time-stamped audit trails for actions that create, modify or delete electronic records.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10e				
Fit Criterion:	An audit trail is generated when creating, modifying or deleting a record.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	66	Requirement Type:	17	Event/use case #:	N/A
Description:	Record changes must not obscure previously recorded information.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10e				
Fit Criterion:	Original file remains intact with the modified file listed as a different record.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	67	Requirement Type:	17	Event/use case #:	N/A
Description:	Audit trails must be retained at least as long as the records to which they pertain.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10e				
Fit Criterion:	An audit trail for a specific record is accessible at any time during the retention of that record.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	68	Requirement Type:	17	Event/use case #:	N/A
Description:	Audit trails must also be available for review and copying.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10e				
Fit Criterion:	The system produces an audit trail on the screen that can be read by the user. Audit trail is captured and printed to the local output device and contains the same information as the recalled audit trail on the screen.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created February 15, 2006				

Requirement #:	69	Requirement Type:	17	Event/use case #:	N/A
Description:	Audit trail will consist of: Image of old data – if modifying or deleting Who created the data When the data was created (time and date stamp) Who modified the data When the data was modified (time and date stamp). Reason for change – if modifying or deleting				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10e				
Fit Criterion:	System creates an audit trail that contains who created the data and when the data was created (time and date stamp), an image of old data, who modified the data and when the data was modified (time and date stamp).				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	

Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created March 6, 2006				

Requirement #:	70	Requirement Type:	17	Event/use case #:	N/A
Description:	System checks must be used to enforce permitted sequencing of steps and events.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10f				
Fit Criterion:	The system has a security check step in place to ensure that only an authorized user is completing a particular step.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	
Dependencies:	N/A	Conflicts:		N/A	
Supporting Materials:	CFR Title 21, Part 11				
History:	Created March 6, 2006				

Requirement #:	71	Requirement Type:	17	Event/use case #:	N/A
Description:	Authority checks must be in place to ensure that only authorized individuals can use the system, access the operation or computer system input or output device, alter a record, or perform an operation.				
Rationale:	Federal Regulation				
Source:	CFR Title 21, Part 11.10g				
Fit Criterion:	The system allows the user to log into the SDMS as an administrator client based on their username and password. The system allows only the specified user to access operations such as print and files captures, alter records or perform other operation.				
Customer Satisfaction:	TBD	Customer Dissatisfaction:		TBD	

Dependencies:	N/A	Conflicts:	N/A
Supporting Materials:	CFR Title 21, Part 11		
History:	Created March 6, 2006		

17b. Standards requirements

See step 17a for all compliance and cGMP/cGLP standards requirements.

18 Open Issues

Issues that have been raised and do not yet have a conclusion.

No known issues at this time.

19 Off-the-Shelf Solutions

19a. Is there a ready-made product that could be bought?

The following COTS solutions are available for purchase if we so choose to purchase a SDMS:

Vendor	**Product Name**
Agilent	Cerity ECM
Amartus	TargetWatch
Computer Compliance Inc.	E-Flexion
Waters	NuGenesis SDMS
Biotrue	Biotrue CDMS
Abrevity	FileData Manager

19b. Can ready-made components be used for this product?

Not applicable.

19c. Is there something that we could copy?

Not applicable.

20 New Problems

20a. What problems could the new product cause in the current environment?

No known issues at this time.

20b. Will the new development affect any of the installed system?

No

20c. Will any of our existing users be adversely affected by the new development?

No

20d. What limitations exist in the anticipated implementation environment that may inhibit the new product?

None

20e. Will the new product create other problems?

No

21 Tasks

21a. What steps have to be taken to deliver the system?

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

21b. Development phases

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

22 Cutover

22a. What special requirements do we have to get the existing data, and procedures to work for the new system?

This section does not apply to this requirements specification. There is no existing data to cutover from an existing system.

22b. What data has to be modified/translated for the new system?

This section does not apply to this requirements specification. There is no existing data to cutover from an existing system.

23 Risks

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

24 Costs

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

25 User Documentation and Training

This section is out of the scope of the purposes of this requirements specification for this specific project; however, in a real setting to deploy a solution to a quality control laboratory at Eli Lilly, a global training team (for the product) would develop and administer training to the end-users. Local OJT for items specific to a laboratory would be administered by a local subject matter expert or team of experts.

26 Waiting Room

This section is out of the scope of this set of requirements, because the requirements are being created "post" product purchase in order to show requirements gathering for a SDMS.

27 Ideas for Solutions

There are none at the present time.

Appendix C- Vendor Comparison Matrix

Feature Available: Y = Yes; N = No; UD = Undetermined; P = Partially; CNT = Could not test

*Requirement No.	NuGenesis SDMS	Cerity ECM	TargetWatch	E-Flexion	Biotrue CDMS	Abrevity FDM
1	Y	Y	Y	Y	Y	Y
2	Y	P	Y	Y	Y	Y
3	Y	Y	Y	Y	Y	Y
4	Y	Y	Y	UD	Y	UD
5	Y	Y	Y	Y	Y	Y
6	Y	Y	Y	Y	Y	Y
7	Y	Y	Y	Y	Y	UD
8	Y	Y	Y	Y	Y	Y
9	Y	Y	UD	UD	UD	UD
10	Y	Y	Y	Y	P	Y
11	Y	Y	Y	Y	P	UD
12	Y	Y	Y	Y	UD	UD
13	Y	Y	Y	Y	UD	UD
14	Y	Y	UD	UD	UD	UD
15	Y	Y	Y	Y	UD	UD
16	Y	UD	Y	Y	UD	UD
17	Y	Y	Y	Y	UD	UD
18	Y	Y	Y	Y	UD	UD
19	Y	Y	UD	UD	UD	UD
20	Y	Y	Y	Y	Y	Y
21	Y	Y	Y	Y	Y	Y
22	Y	Y	Y	Y	Y	Y
23	Y	Y	Y	Y	Y	Y
24	Y	Y	Y	Y	Y	Y
25	Y	Y	Y	Y	Y	Y
26	Y	Y	Y	Y	Y	Y
27	Y	Y	Y	Y	UD	UD
28	Y	Y	UD	UD	UD	UD
29	Y	Y	Y	Y	UD	UD
30	Y	Y	Y	Y	Y	Y
31	Y	Y	Y	Y	Y	Y

*Requirement No.	NuGenesis SDMS	Cerity ECM	TargetWatch	E-Flexion	Biotrue CDMS	Abrevity FDM
32	Y	Y	Y	Y	Y	Y
33	Y	Y	UD	UD	UD	UD
34	UD	UD	UD	UD	UD	UD
35	UD	UD	Y	Y	Y	Y
36	Y	Y	Y	Y	Y	Y
37	Y	Y	Y	UD	UD	Y
38	Y	Y	CNT	CNT	CNT	CNT
39	Y	Y	CNT	CNT	CNT	CNT
40	Y	Y	CNT	CNT	CNT	CNT
41	Y	Y	CNT	CNT	CNT	CNT
42	Y	Y	UD	UD	UD	P
43	CNT	CNT	CNT	CNT	CNT	CNT
44	CNT	CNT	CNT	CNT	CNT	CNT
45	CNT	CNT	CNT	CNT	CNT	CNT
46	CNT	CNT	CNT	CNT	CNT	CNT
47	CNT	CNT	CNT	CNT	CNT	CNT
48	CNT	CNT	CNT	CNT	CNT	CNT
49	CNT	CNT	CNT	CNT	CNT	CNT
50	CNT	CNT	CNT	CNT	CNT	CNT
51	CNT	CNT	CNT	CNT	CNT	CNT
52	Y	Y	Y	Y	Y	Y
53	Y	Y	Y	Y	Y	Y
54	Y	Y	Y	Y	Y	Y
55	Y	Y	Y	Y	Y	Y
56	Y	Y	Y	Y	Y	N
57	Y	Y	Y	Y	UD	N
58	Y	Y	Y	Y	UD	N
59	Y	Y	Y	Y	Y	N
60	Y	Y	Y	Y	UD	N
61	Y	Y	Y	Y	Y	N
62	Y	Y	Y	Y	Y	N
63	Y	Y	Y	Y	Y	N
64	Y	Y	Y	Y	UD	UD
65	Y	Y	Y	Y	Y	Y

*Requirement No.	NuGenesis SDMS	Cerity ECM	TargetWatch	E-Flexion	Biotrue CDMS	Abrevity FDM
66	Y	Y	Y	Y	*Y	Y
67	Y	Y	Y	Y	*Y	Y
68	Y	Y	Y	Y	*Y	Y
69	Y	Y	Y	Y	*Y	Y
70	Y	Y	Y	Y	*Y	Y
71	Y	Y	Y	Y	*Y	Y

Number Reference	
1-5:	Platform
6-11:	Native Instrument Data Archived
12-19:	Data
20-22:	Metadata
23-32:	Processed Data Migrated
33-34:	Data
35-37:	Look & Feel
38-42:	Usability & Humanity
43-47:	Performance
48-49:	Operational
50-51:	Maintainability & Support
52-61:	Security
62-71:	Legal

*Y: Vendor stated that is it supported; however, no supporting documentation was provided.

Note: Some inferences were made based on the general concept of a requirement section

Appendix D: User Questionnaire

The following questionnaire has been developed to identify which SDMS are being used in pharmaceutical laboratories, and to gain a general knowledge of how they are being used. In order to protect people, as well as the company's current business interests, no company names will be used.

Please provide the following information.

Date of Questionnaire completion: See below

Title: See below

	Title	Date Received
Response 1:	Sr. Systems Analyst – Analytical Sciences R&D-IT	01/19/2007
Response 2:	Sr. Systems Analyst – Analytical Sciences R&D-IT	02/02/2007

Questions:

1. Which SDMS do you currently use (if any)?

Q1	
Response 1:	7.0 SR-1
Response 2:	Waters NuGenesis SDMS

2. For what reasons did you choose this SDMS over others on the market?

Q2	
Response 1:	Previous experience in chemistry labs showed good productivity increases. Also implemented as a solution to address quality concerns of the lack of control over raw data in the laboratories.
Response 2:	▪ Reports stored in scalable vector graphic format ▪ Report content and metadata indexed and searchable ▪ Robust, stable archive engines ▪ Ability to use with a vast variety of instruments

3. How is the SDMS integrated into your laboratory's workflow?

Q3	
Response 1:	SDMS is responsible for capturing printed reports. Some printed reports – instrument outputs – are linked and made available to LIMS sample submitters. The printed lab outputs are also available for inclusion in the e-Lab notebook write-ups.
Response 2:	SDMS is central repository of data files and reports from multiple lab instruments. Files are archived (copied from lab instrument controller into repository) and managed (deleted from local disk) automatically on a preset schedule. Schedule is customized for each instrument. Graphical results/printable reports are entered into SDMS manually by users as needed, at the user's discretion. ▪ Provides compliance and security by securing instrument data. ▪ Frees up disk space on instrument controllers (by deleting older data after archival) ▪ Provides a central searchable repository for storing graphical results ▪ Being integrated with other lab informatics apps to provide access to these graphical reports

4. Has this SDMS improved efficiency? If so, how? If not, why?

    - Improving scientist collaboration via electronically shared data?
    - Streamlining access to data sources?
    - Automating delivery of new information?

Q4	
Response 1:	Yes. Allows for the implementation of a 'paperless' lab workflow environment. Reduces long term costs for archival management of paper artifacts. Allows for quick retrieval of relevant files or printed data. Yes. Yes. Holds the instrument outputs and raw data. Not really. People consume the contents. No automation, like verification, is currently implemented.
Response 2:	Yes. See answer to #3.

5. If anything, what would you have or like to do differently about how the SDMS was integrated into your workflow?

Q5	
Response 1:	Nothing that I can think of. We have in place what is allowed by the various API's.
Response 2:	Each instrument uses a different report format. Even similar instruments (e.g. FTIR) using same software (e.g. OMNIC v6.0) generate different reports. So no consistent fail-safe way of harvesting typical metadata. So a lot of metadata for a report must be filled out manually by the user when entering the report into SDMS. On data file side, there are no component-wide file/folder naming conventions. So a quick, standard way of harvesting metadata is not readily available. If time, resources and user support were available, I would have liked to standardize/ normalize reports formats, file naming and storage conditions across labs before the application was rolled out.

6. (If you answered ‘No’ to Question 4): Do you think there is a better SDMS on the market that would have better fit your needs?

Q6	
Response 1:	Other products have better support for PDF documents and some data transformation capabilities into XML. Not a show stopper at this time, but Waters SDMS has areas that need to be improved.
Response 2:	N/A

VITA

JOSEPH EDWARD HEYWARD II

jheyward2@yahoo.com
7642 Cherryberry Drive
Indianapolis, Indiana 46239

EDUCATION

Indiana University-IUPUI — Indianapolis, IN
M.S. in Laboratory Informatics — 2008

Purdue University-IUPUI — Indianapolis, IN
A.S. in Computer Technology — 2004

Additional coursework in Organizational Leadership and Supervision
(Applied Leadership and Leadership Philosophy)

Hampton University — Hampton, VA
B.A. in Chemistry — 1994

AWARDS/HONORS

Who's Who among Students in American Universities, 2003
Tau Alpha Pi Honor Society – Phi Beta Chapter, 2002

TECHNICAL PROFICIENCIES

Platforms: Windows 95/98/NT/Millenium/2000/XP
Laboratory Instrumentation: HPLC, GC, TLC, UV/VIS, IR, AA, FTIR, Karl Fischer
Laboratory Software: Waters Empower, Agilent Chemstation, JMP, NovaManage, Lab Materials Inventory (LMI), Finish Results Control Charting (FRCC), LabView, Labware LIMS, Darwin LIMS
Computer Software Tools: Microsoft Office Suite, Adobe Acrobat, Documentum, Remedy, SQL Dreamweaver, Trackwise
Computer Programming Tools: Visual Basic, HTML, and coursework in C++ and UML

CONFERENCES ATTENDED

Laboratory informatics: Transforming Analytical Data to Information to Knowledge, 2006

PROFESSIONAL EXPERIENCE

March 2007- Present Eli Lilly and Company Indianapolis, Indiana

Senior QA Associate - Business Integrator for Global Quality LIMS Deployment

- Subject matter expert for quality control and sample management processes.
- Work with site resources to determine proper system configuration and identify the impact of new business processes and IT systems in their area.
- Advocate for the site users for future requirements gathering and implementation.
- Drive business process change where necessary to gain maximum benefit from the solution.
- Influence sites to align local business practice with recommended global business processes.
- Communicate with business users to understand local requirements and apply these to global business processes and/or system configuration.
- Perform acceptance testing on the various software builds.

September 2006 – March 2007 Eli Lilly and Company Manassas, Virginia

QC Representative - Laboratory Informatics

Provide site support for QC informatics solutions by:

- Implemented and maintained various laboratory solutions including LIMS, Chromatography, Environmental Monitoring, Stability, Instrument Maintenance, and some stand-alone solutions.
- Applied expertise in the control strategies and implementation of these solutions for Chemistry, Microbiology, Devices, and Inbound Testing laboratories.
- Interfaced effectively with site validation, lab metrology, Computer Systems Quality, Manufacturing Informatics & Computer Systems, and global business partners.
- Provided direction in developing a paper-light business environment for laboratory operations.
- Wrote standard operating procedures and computer system validation documents.

October 2000 – August 2006 Eli Lilly and Company Indianapolis, Indiana

Laboratory Informatics Chromatography Data System Specialist (2003-2006)

Local Project Coordinator

- Managed resources for validation and qualification.
- Coordinated local implementation of the CDS with the Global team.
- Ensured all deployment tasks are completed on time
- Promoted issues to appropriate resources as needed

- Designed a Microsoft Access database to track the progress of the method conversion from the Lilly Chromatography system (LCS) to the Waters Millennium CDS/GCDS.

Global Change Control Board Representative

- Provided high-level user requirements and overall business knowledge for requirements gathering and deployment impact assessment
- Single point of contact for my plant site and department
- Kept plant site and department informed of all GCCB proceedings.
- Evaluated and approved and rejected global change requests.
- Participated in global client acceptance testing.

Lab Data Owner and Power User

- Approved and revoked security access the CDS laboratory data projects.
- Verified training prior to account request approvals.
- Reviewed and approved specific local CDS documents (e.g. System Requests, validation documents).
- Wrote local change control requests and perform the required action steps for the implementation of new system releases.
- Wrote and executed data verification protocols for the conversion of the methods from the Lilly Chromatography System (LCS) to the Waters Millennium CDS/GCDS.
- Provided local configuration support and method management.

Senior Analytical Chemist, Diversified Hospital Care (DHC) (2000-2003)

- Provided support in areas of technical troubleshooting, analytical testing and second-person verification and non-analytical work, such as, performing method transfers, reviewing and approving technical documents, and writing out-of-specification (OOS) and deviation reports.
- Performed investigations and root cause analysis of atypical and out-of-specification results.

March 1995- October 2000 Procter and Gamble Greensboro, North Carolina

Analytical Chemist, Personal Health Care (PHC)

- Performed USP/NF compendia analysis on raw materials and tested finish product per validated and verified methods.
- Designed and implemented IQ/OQ validation protocols for laboratory equipment.
- Reviewed and revised SOP's to comply with corporate and FDA guidelines.
- Delivered policy and procedural training to other lab analysts, including USP/NF and finish product methods and documentation requirements as defined by FDA and GMP regulations.
- Interviewed potential new employees.

- Served as the capital budget manger for laboratory expansion project.
- Appointed the position of the out-of-specification coordinator.

Out-of Specification (OOS) Coordinator

- Assessed OOS data to ascertain if results may be attributed to laboratory error, instrument malfunctions or bad product from suppliers or misformulated during the making process.
- Conducted OOS investigations and wrote the final root cause analysis reports.
- Applied corrective action plans and systemic corrections by targeted delivery dates.
- Provided periodic updates of investigations to management, which included trend analysis for identifying training issues and opportunities for improvement.

Lab Design & Equipment Owner for Pepto-Bismol Project

- Managed $6000M laboratory project budget and saved $200M in expansion costs.
- Designed the layout for the laboratory expansion.
- Purchased analytical equipment for Bismuth Subsalicylate (BSS) and Pepto-Bismol testing.
- Coordinated schedules and directed vendors and construction crews during lab expansion.
- Installed and inspected all gas lines that connected the HPLC and GC instruments.

PROFESSIONAL AFFILIATIONS

Alpha Phi Alpha Fraternity, Incorporated
Big Brothers/Big Sisters Volunteer
Association of Computing Machinery (Former Member)
American Chemical Society (Former Member)