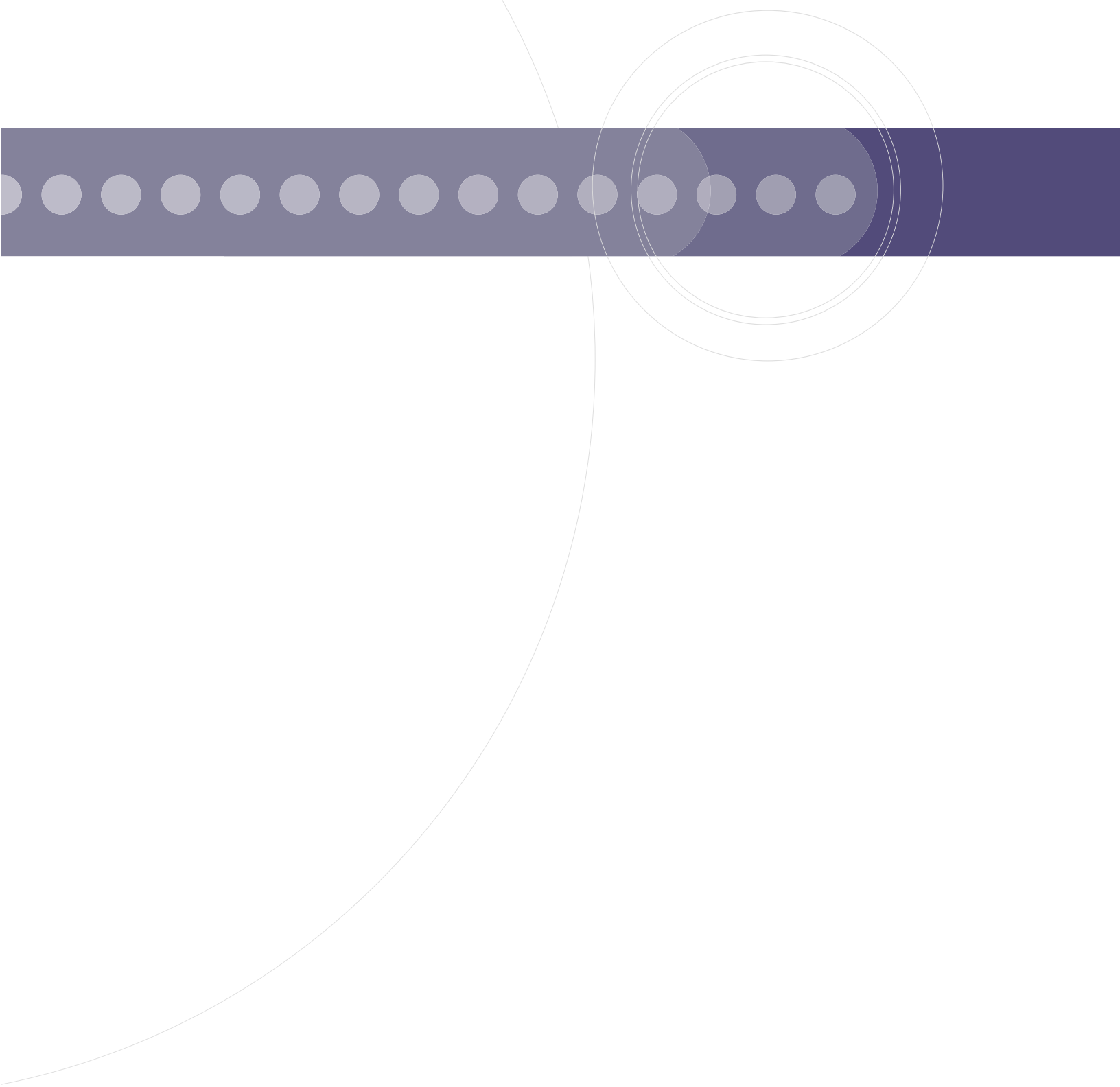


Regulatory Compliance in Life Sciences

Leveraging Integration Competency Centers to Manage Information Risk





While every attempt has been made to ensure that the information in this document is accurate and complete, some typographical errors or technical inaccuracies may exist. Informatica does not accept responsibility for any kind of loss resulting from the use of information contained in this document. The information contained in this document is subject to change without notice.

The incorporation of the product attributes discussed in these materials into any release or upgrade of any Informatica software product—as well as the timing of any such release or upgrade—is at the sole discretion of Informatica.

This edition published November 2004

Table of Contents

Executive Summary	2
What Does Compliance Mean?	3
Why Comply?	3
Proving Compliance	4
How to Comply	4
Understand the Nature and Content of Data	4
Establish a Single Source of Truth	4
Access It Rapidly	5
Benefits of Integration	5
Platform for Compliance	6



Executive Summary

Despite a general slackening in the growth rate of the world economy, the life sciences sector—biotechnology, pharmaceuticals, diagnostics, academic work, and agriculture—is growing at a compound annual rate of nearly twice that of any other sector. For these companies, intellectual property is their most valuable asset—and the potential for mishandling it is their greatest risk.

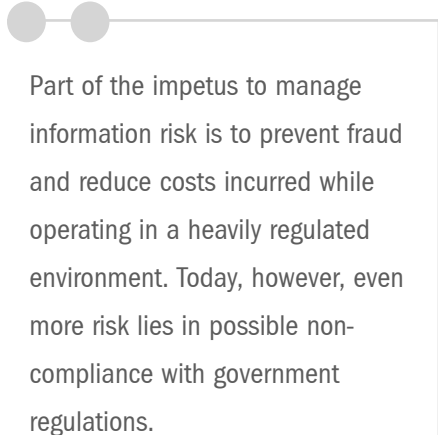
This white paper describes information risk management within life sciences enterprises. Managing information risk calls for organizations not just to understand the nature and content of their data, but also to be able to prove its accuracy and access it rapidly.

Part of the impetus to manage information risk is to prevent fraud and reduce costs incurred while operating in a heavily regulated environment. Today, however, even more risk lies in possible non-compliance with government regulations. New and existing legislation—such as the Data Protection Acts in Europe, Sarbanes Oxley in the US, and a raft of industry regulations, such as those imposed by the Federal Drug Administration and its EU counterparts—cover diverse issues ranging from freedom of information to the dumping of hazardous waste. Any business today can be called to account at a moment's notice, often for transgressing laws of which it was not even aware.

Compliance requirements affect business processes, many legacy systems, and closed working environments. Of those companies surveyed¹, 40 percent said their biggest challenge was bringing legacy systems into compliance. Many companies said that simply interpreting compliance regulation was one of their greatest challenges. Unfortunately for IT departments, more than half of the senior managers in the life sciences sector view the role of delivering compliance as being a strictly IT problem.²

In the same study, fewer than 10 percent of respondents in the major pharmaceutical organizations consider themselves to be fully compliant on all systems. The rest hope they don't get caught. Many perceive paying a fine to be less expensive than implementing a new system. However, any reactive implementation is bound to be far more costly and less versatile than a planned, prepared, and fully scoped project.

This paper examines some of the issues around implementing technologies to increase compliance in the life sciences sector. It discusses what compliance means, why companies must comply (and prove compliance), and then provides recommendations on how to comply, by understanding the nature and content of data, establishing a single source of truth, and being able to access it. It also recommends creating or leveraging integration competency centers (ICCs) to bring a company's systems into compliance, and to deliver the efficiency and productivity benefits of a services oriented architecture (SOA).



Part of the impetus to manage information risk is to prevent fraud and reduce costs incurred while operating in a heavily regulated environment. Today, however, even more risk lies in possible non-compliance with government regulations.

¹ Gartner G2 survey of pharmaceuticals May 2003
² Gartner G2 survey of pharmaceuticals May 2003

What Does Compliance Mean?

Compliance with currently accepted procedures helps protect a company from risks that come from many different directions. Dissatisfied patients, customers, suppliers, or staff might take action against a company for some perceived misdemeanor. The large and growing volume of government legislation is an even more serious threat. In addition to more established legislation, companies need to address the implications of FDA 21 CFR Part 11, Sarbanes Oxley, the USA PATRIOT Act, and EU laws covering freedom of information and data protection. All these acts demand that organizations not only prove the accuracy and authenticity of their data, but in many cases also how it was derived and on what information conclusions were based. For companies regulated overseas or with a US or foreign parent or correspondent company, the burden is even greater.

In addition to the legislative burden imposed by these various acts, organizations are also required to be able to readily lay their hands upon a wealth of information to satisfy other bodies. They may be called upon to produce information that documents the submission of tax returns, personnel management procedures, health and safety compliance, access for the disabled, and virtually anything else that is regulated in any way.

Each set of regulations imposes its own requirements for the retention and analysis of data, but compliance with one does not mean compliance with all. For example, while common business practice may be that emails are held for one year, taxation authorities can demand data going back six. From a litigation perspective, the general rule should be that any information that could be pertinent should be retained for as long as the possibility of litigation exists.

Why Comply?


The risks of non-compliance are high. They include loss of credibility, reduced shareholder confidence, and significant adverse publicity. Even more serious for the individuals involved is the prospect of being fired, with little chance of re-employment, or in the worst case, imprisoned.

The key regulation of concern today for pharmaceuticals is the US Food and Drug Administration (FDA) 21 Code of Federal Regulations (CFR), Part 11. This regulation deals with the criteria under which the FDA considers electronic records, electronic signatures, and handwritten signatures executed to electronic records to be trustworthy. Pharmaceutical companies have limited time to ensure 21 CFR Part 11 compliance. Final guidance for 21 CFR Part 11 has been issued and Gartner Research suggests that by 2007 all FDA regulated industries will have to comply with Part 11 requirements. Non-compliance with Part 11 is no longer an option for pharmaceutical companies.

Warning letters and 483's are costly to address and lead to non-compliance if not successfully resolved. The FDA makes a case-by-case evaluation as to whether to pursue regulatory action against companies that are viewed as out-of-compliance with Part 11. Such actions may include increased scrutiny from regulatory bodies, and can result in costly remediation work and downtime, compromised product quality, criminal court sanctions such as fines, personal liability of directors, managers or other responsible officers of an organization, and even prison sentences.

Compliance: How Does Your Organization Measure Up?

- Do you know how to access the audit trail of creation, amendment and ultimately disposal or retention of each document you have stored? Do you have an appropriate document retention policy in place?
- How much does storage of electronic documents cost your organization? What proportion of the documents stored is actually required to be retained for legal or commercial reasons? What proportion of the documents is actually unnecessary duplication?
- Could your organization respond to an information request from an individual for all documents including e-mails that you hold about them, as required under the UK's DPA?
- Could your organization answer a request from the US FDA for a full breakdown of all activity on a particular drug discovery or trial within ?? days as required ?
- Would your organization be able to respond properly to a discovery request for documents to be submitted to a court in relation to a dispute, ensuring that only the required documents are disclosed and no others? How much would it cost your business?
- Does your organization have an appropriate information request policy in place?
- Would your organization be able to prove the authenticity and integrity of a retained e-mail vital to proof of contractual terms entered into? Are your IT systems able to provide this vital evidential weight?
- How does your organization ensure data is destroyed? Is the data overwritten adequately? Are hard-drives cleaned upon return to a vendor, or physically destroyed? Does your organization have an appropriate IT Hardware Disposal Policy in place?



Non-compliance could even lead to the invalidation of an entire clinical trial. All progress would be halted pending a complete re-collection of all the data—a stunning blow to an application for regulatory approval. Failure to retain appropriate documentation may also lead to uncertainty regarding agreed contractual terms, or difficulties with responding to information requests under the DPA, or orders for discovery in court. Above all, any publicized failure to comply puts organizations’ reputations and commercial standing at risk.

Proving Compliance

Companies need not only to follow best practice in complying with regulations, but also be able to prove they are doing so. Legal cases against a company—whether brought by an aggrieved former employee in a civil suit or by a government body—are won or lost on the weight of evidence. Organizations that can readily access all documents that are relevant to an action, and prove their validity, are likely to see an action dismissed or settled at a far earlier and less damaging stage of the proceedings.

It’s always harder to prove compliance than it is for someone to accuse you of not complying. Much depends on the opinion of the relevant regulator or inspector. The more easily you can supply required information, the better your chance of being viewed as a compliant business. However, few organizations even know the content of 80% of their data. This makes it virtually impossible for them to produce requested data, let alone show where it came from in order to document its probity. One compliance manager comments, “Too many of these guys (the regulators) have been brought up to believe that we run systems like those in Star Trek where you can press a button and instantly access the total knowledge of the entire universe. The reality is that it can take days, or even weeks, to find the data they want.”

How to Comply

Understand the Nature and Content of Data

Life science industries deal with many different types of data, from knowledge management and database systems through specialized applications for informatics and sequencing, proteomics (protein research) through computational chemistry to data visualization. Much of the data comes from sources outside the enterprise, and is generally found in heterogeneous, often proprietary formats that stand in the way of easy data sharing. The data sources could be structured, semi-structured, unstructured, in relational databases or flat files.

Despite that complexity, there are a number of technology solutions that help to identify the content of a data file and describe it either by the file type, e.g. email, graphic or text. There are also a number of more sophisticated applications that can also describe its content. Data about data is called metadata, and it’s crucial to effective management of data. Metadata can tell you what type of information is in a document, whether it has been altered and by whom. It also enables automatic classification of files for more rapid and effective data searches in support of your organization’s storage management policy. Comprehensive metadata tagging enables an organization to understand the nature and type of the data it is storing, where it comes from, and how it has been used.

There are two distinct types of metadata: business metadata and technical metadata. The role of business metadata is to ensure that all results are returned within the correct context. For example, it identifies data as regulatory, financial, or specific to a particular subject such as a medication or a phase of a trial. It also records who has read/write access to the data and who has actually made any changes, and when. Business metadata helps organizations identify which

data elements are the same even though they may be implemented in different systems and look different. Business metadata enables the business to leverage information management resources to capture, integrate, and publish data from diverse sources to support multidisciplinary teams. It enables safer and faster throughput of data, whether for clinical trials or for marketing and sales of new drugs.

Technical metadata focuses on physical implementations of data and systems data such as passwords, backup locations, and timing. Technical metadata provides an audit trail of who queried the data and when, and provides a lineage of the processes were conducted against the data, including transformations, cleaning operations, and corrections.

Establish a Single Source of Truth

In response to having data in diverse sources across the enterprise, many organizations have chosen to extract a copy of the information held in the company’s operational IT systems and pool it into a data warehouse. Once the data warehouse has been created, all the different types of information are held in a common format, regardless of their source. This common bank of data can then be used as a source for any business intelligence or analytic application, and allows the company to achieve the prime objective of business intelligence—a single trustworthy source of information.

Companies can enhance data warehouses so that they track the history surrounding each record: identifying who changed the data, showing prior and current values, detailing when it was changed, and even recording electronic signatures. If companies add metadata analysis functionality, the data warehouse can also assist with audits, such as establishing who queried an item of data and when, as well as provide traceability to determine what processes were conducted against it such as transformations, cleaning operations, and corrections. A data

warehouse also offers the ability to conduct impact analysis across multiple systems with multiple standards.

Certain data objects are key in order to demonstrate compliance with 21 CFR Part 11. These cover the information about what data was changed; who changed it; when it was changed; current and prior values and the reason for change. If an organization has an appropriate underlying metadata repository, it can generate a report about any attribute for which data was collected. For example, clinical attributes such as a subject's date of birth, his or her weight at the time of screening, together with the date and time that treatment was given, could be collected in addition to the clinical data that is required for 21 CFR Part 11 compliance. Such a report allows managers to demonstrate that they can show who has changed the data collected for a clinical trial.

Access It Rapidly

Leveraging the properties or attributes of data objects and related elements allows information to be stored and then rapidly accessed in a meaningful format. The retrieval process can deliver a set of reports to provide organizations with insight about the extent to which they are compliant with Part 11. Those reports could be highly visual: dashboards or charts that make the information even more readily accessible. The reports might also allow for real-time updates from queues and streaming data, bypassing even the data warehouse.

Whenever a company's needs change, and it wants to integrate and access data in a different way, it shouldn't have to start from scratch. Reuse enables a company to benefit from its previous investment in design, development, and testing. Companies need the capability to tweak or re-deploy existing architectures, data services, and objects to serve new purposes, without having to invest significant development time or resources.

Benefits of Integration

To understand the nature of all an organization's information, to establish a single source of truth, and to be able to access it rapidly, life sciences companies need to integrate their disparate systems. Proving compliance with regulations is virtually impossible without integration. One company interviewed in the research for this paper pointed out that under FDA requirements it could be asked to produce a complete picture of the research, manufacture, trials, delivery and usage of one product information—and that, to respond to that request, it would have to draw information from 26 IT systems. As each system has its own unique way of identifying and describing its records, and is in a different vendor's database application, answering this single request could tie up an entire IT department for weeks.

Integration has business benefits that extend beyond compliance, too. Better data security and knowledge-sharing lead to greater efficiency in clinical trials. For example, a company could establish standard performance measurement scorecards for recording and analyzing the time and money spent on each stage of trials. This could be enhanced with time-based metrics to track trial duration from milestone to milestone, enabling faster submission, faster approvals, and overall cost savings.

According to Gartner's analysts, nearly half of all pharmaceutical companies expect their IT departments to take ownership of the compliance issues and solve them single-handed. While some software vendors would like to suggest that the answer to these problems lies in adopting a single-vendor environment—theirs—that would be a prohibitively expensive and potentially inflexible solution for most companies. There's no one-size-fits-all solution to the challenges of bringing together all the different kinds of functionality and information that a life sciences company requires.

What's an Integration Competency Center?


In contrast to ad-hoc integration projects conducted through different business units, an ICC is a shared IT function that enables project teams to complete a data integration effort rapidly and efficiently by following best-practice processes, leveraging the expertise of staff with integration-specific roles, and utilizing standard technologies. According to Gartner's Diane Morello, "[Competency centers] often span time zones, continents and corporate entities. They identify and build best practices in areas of expertise, and they help deploy people effectively, whether internal staff, external staff or a blend."³

Key objectives of an ICC:

- Lead and support integration projects with the cooperation of subject matter experts
- Promote data integration as a formal discipline
- Develop staff specialists in integration processes and operations and leverage their expertise company-wide
- Assess and select integration technology and tools
- Manage integration pilots

Today's best ICCs leverage service oriented architectures (SOA) to create a modular set of universal data services (UDS). These deliver a consistent approach to defining and accessing data sources, enabling organizations to streamline their data integration architecture while enabling future flexibility and adaptability.

³ 18 July 2002; Morellos, Diane; Back to Basics: What are Competency Centers



Why Companies Choose Informatica for ICC:

- Live/shared view of the entire production environment
- Maximized reuse (systems, processes, resources, and interfaces)
- Optimized performance (known dependencies and assets)
- Robust repository (embedded rules and relationships, sharing model)
- Service-oriented architecture and support for universal data services

Instead, the most effective approach is to consider all the systems in use within the organization and identify the opportunities to create a unified integration architecture that brings together all relevant data and puts it into a common and widely usable format. By anchoring the architecture on a common data server and creating a service-oriented architecture (SOA), IT can easily provide shared services for data access, integration, audit, and visualization. These services are critical for a wide variety of integration solutions, whether built by corporate IT or by third-party partners. This is especially important as end users grow to recognize the importance of an audit trail for proving compliance.

There's no single blueprint for creating an SOA: much depends on the information assets a company already has and how it chooses to leverage them. The process of implementing an SOA may include creating a data warehouse, migrating, consolidating, and synchronizing systems.

Integrating data in a data warehouse requires an organization to invest some time and resources up-front in developing and implementing a methodology. However, once the data warehouse has been set up for one data source, such as discovery records, as much as 80 percent of that effort can be re-used on subsequent projects, making it much faster and less resource-intensive. Data warehouses reduce the need for batch data processing—one of the industry's biggest IT headaches—and reduce the need to store multiple copies of similar data in different forms for different applications. Data warehousing also helps overcome one of the biggest challenges facing life sciences organizations: the lack of a single, accurate picture of all their activities.

Platform for Compliance

The most flexible and efficient approach to solving the challenges of compliance is to track metadata and integrate information from multiple systems in an SOA. Flexible, easily deployed, and smart shared data services help to eliminate data silos and simplify integration efforts for consistent views of information across the enterprise.

Underlying an SOA is a ubiquitous integration layer across the entire enterprise architecture, to support the dynamic assembly of data services. This data integration platform is fast becoming a requirement for keeping pace with growing data demands and infrastructure complexity, particularly in the light of ever-more demanding compliance requirements.

A data integration platform enables shared data services for access, integration, auditing, and visualization to come together on an as-needed basis. For a life sciences organization, not only does it allow the business to harness and validate legacy data, it also positions it to move forward with the adoption of new platforms and technologies, such as Web-based systems for clinical trials development and discovery. These systems are expected to account for more than 50 percent of all R&D activity by the end of 2005, and will be ubiquitous within a few years.

One key attraction of a data integration platform is its emphasis on re-use of routines and operations. This differs from code-based or code-generating solutions that rely heavily on developer expertise, take more time to develop, are error prone, and hard to modify and reuse. Instead, data integration platforms use an object-oriented approach that is fundamentally codeless, separating logical data integration design from the physical operational environment, allowing designers to focus on the "what" and not the "how." The result is increased flexibility to develop integration solutions quickly and easily evolve them to adjust to change—including changes in regulations that require a company to alter its business processes.



INFORMATICA®

Worldwide Headquarters, 2100 Seaport Boulevard, Redwood City, CA 94063, USA
phone: 650.385.5000 fax: 650.385.5500 toll-free in the US: 1.800.970.1179 www.informatica.com

Informatica Offices Around The Globe: Australia • Belgium • Canada • France • Germany • Japan • the Netherlands • Singapore • Switzerland • United Kingdom • USA

© 2004 Informatica Corporation. All rights reserved. Printed in the U.S.A. Informatica, the Informatica logo, Turning integration into insight, Informatica PowerAnalyzer, PowerCenter, Informatica SuperGlue, are trademarks or registered trademarks of Informatica Corporation in the United States and in jurisdictions throughout the world. All other company and product names may be tradenames or trademarks of their respective owners.

J50287 6545 (12/08/04)