Delivering information you can trust

October 2008

Information Management software

Accelerating an information agenda with IBM InfoSphere Foundation Tools

Contents

Taking the first step toward an information agenda

Organizations all over the world have adopted information-intensive initiatives to improve operational efficiency, reduce costs and gain competitive advantage. A wide range of projects, including master data management (MDM), data warehousing, business performance management and enterprise content management, are helping businesses improve their core competencies and achieve point objectives.

Many organizations are ready to advance even further down the road of information-based innovation. Executives and management teams want to understand and monitor critical key performance indicators (KPIs) for the organization to make the right business decisions and better manage the bottom line. Analysts and business users are interested in monitoring trends in information and using that data to modify procedures accordingly.

To take this next step, businesses must streamline and connect information and systems across enterprise domains with an integrated information infrastructure. The point solutions deployed to date are built to handle specific tasks, but often do not take into account the need to connect across enterprise departments to achieve true information-based innovation. Critical data may be trapped in departmental silos and inaccessible, or stored in systems that were originally built for speed and performance, not access and analysis.

This disconnected information architecture leaves IT organizations unable to respond rapidly to new information requests from business users and executives. With few tools or resources to track the information sprawl, it is also difficult for businesses to monitor data quality and consistently apply business rules. As a result, information remains scattered across the enterprise under a myriad of disorganized categories and incompatible descriptions.

The IBM® Information Agenda is specifically designed to help organizations address these challenges and build a robust information architecture that leverages existing IT investments. It offers a proven approach to identifying vital information; specifying how, when and where it should be made available; determining data management processes and governance practices; and aligning the use of information to match an organization's business strategy.

A crucial part of adopting an information agenda is gaining a clear understanding of the information that exists within business systems. Organizations must assess their current information landscape to determine what information is available and what information is needed to drive greater business value, and then devise a phased plan for bridging the gap.

IBM InfoSphere™ Foundation Tools are designed to address this need by helping organizations profile, model, define, blueprint and govern their information. With InfoSphere Foundation Tools, organizations can discover and design their information infrastructure and start building trusted information across the enterprise.

The four facets of an information agenda

An information agenda helps organizations transform their information into a trusted strategic asset that can be rapidly leveraged across multiple applications, processes and decisions for sustained competitive advantage. It can help an organization:

- *Identify data and content that are vital to the organization*
- *Identify how, when and where information should be made available*
- *Determine data management processes and governance practices*
- *Identify and prioritize the information projects that deliver the most return*
- *Align the use of information to match the organization's business strategy*
- *Create and deploy an information infrastructure that meets both immediate and future needs*

An information agenda has four key components:

- ***Strategy:*** *A shared, comprehensive vision that helps the organization identify and prioritize business objectives and establishes an information-driven strategy to achieve those objectives.*
- ***Define and govern:*** *This component focuses on discovering and designing trusted information with unified tools and expertise to sustain competitive advantage over time*
- ***Information infrastructure:*** *The foundation that identifies the technology components and capabilities needed to establish a common information framework. The infrastructure is built with open and agile technologies, including information integration, MDM and BI, and leverages existing information assets for speed and flexibility*
- ***Roadmaps:*** *A phased execution plan for transforming the organization, the prescriptive roadmap helps accelerate information-intensive projects that are tightly aligned with your strategy—speeding both short-term and long-term return on investment*

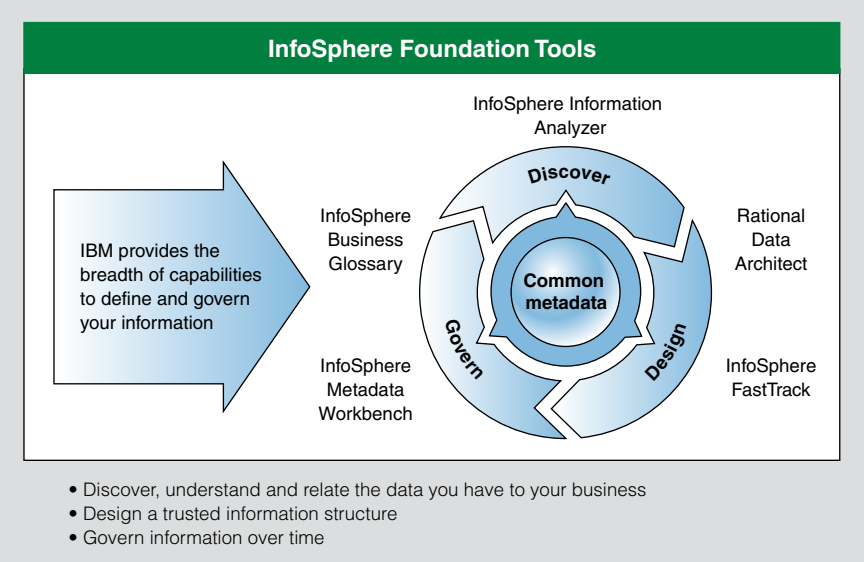
InfoSphere Foundation Tools: An overview

InfoSphere Foundation Tools are an open set of tools that help prepare an organization to adopt an information agenda and work to ensure its success. InfoSphere Foundation Tools include:

- *IBM InfoSphere Information Analyzer*
- *IBM Rational® Data Architect*
- *IBM InfoSphere FastTrack*
- *IBM InfoSphere Metadata Workbench*
- *IBM InfoSphere Business Glossary, including Business Glossary Anywhere*

These five core components can be deployed together or independently depending on the specific needs of the organization (see Figure 1).

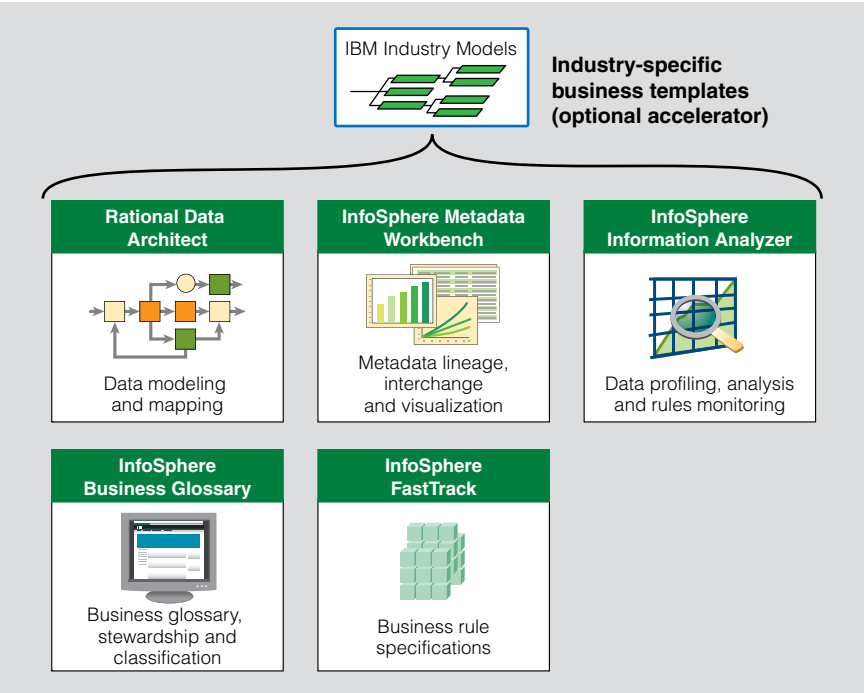
Figure 1: The five components of InfoSphere Foundation Tools help an organization align enterprise information with business goals.



InfoSphere Foundation Tools provide a unique combination of capabilities to help organizations understand and document their existing data resources. These tools are used to discover and document new and existing information sources across the enterprise to better understand their structure, content and quality. The tools can also help organizations understand where a particular piece of information comes from, its usage and its business meaning.

The tools help companies align business and IT objectives by defining common data models and a common business language (see Figure 2). In addition to creating a common business vocabulary across the enterprise, organizations can also inventory where that business concept is represented across physical information structures. Data stewards can be assigned to manage the business language, and enterprise users can collaborate with these stewards to keep the central vocabulary up-to-date and consistent. The dimensions and hierarchies that represent these business constructs can also be defined through the tools.

Figure 2: InfoSphere Foundation Tools address a wide range of business processes, from data modeling and profiling to information governance and accountability.



Accelerating the deployment of InfoSphere Foundation Tools with IBM Industry Models

Businesses can accelerate their information design processes by using the IBM Industry Models in tandem with InfoSphere Foundation Tools. The fully customizable IBM Industry Models supply best-practice data and business process models as well as KPIs based on proven industry examples. Developed over 10 years using IBM's experience with more than 400 clients, the IBM Industry Models support the six major industry verticals: banking, insurance, financial markets, retail, telecommunications and healthcare (see Figure 3).

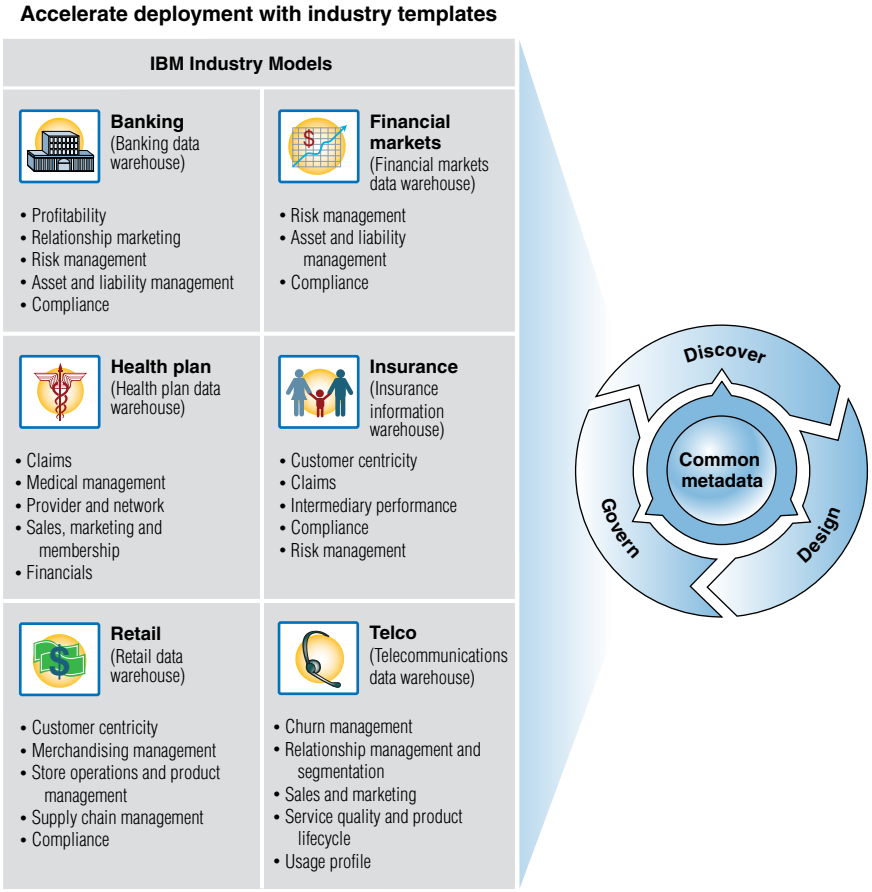


Figure 3: Predefined IBM Industry Models provide industry-specific data and process models to help speed deployment of organizations' information-centric initiatives.

The IBM Industry Models include a glossary of business terms, concepts and a physical data model. These models can be used to prepopulate the InfoSphere Foundation Tools' InfoSphere Business Glossary to share common definitions and the location of assets across the enterprise; help better align business and IT users; and help accelerate overall project delivery. The templates allow organizations to expedite industry-specific business and technical metadata to expedite data integration projects such as MDM initiatives or data warehouse development.

Who are IBM InfoSphere Foundation Tools for?

IBM InfoSphere Foundation Tools were designed for data analysts, business analysts, subject matter experts, architects and governance stewards. These individuals must ensure that the IT information architecture supports all information-centric projects, including MDM, business intelligence applications, governance initiatives and more.

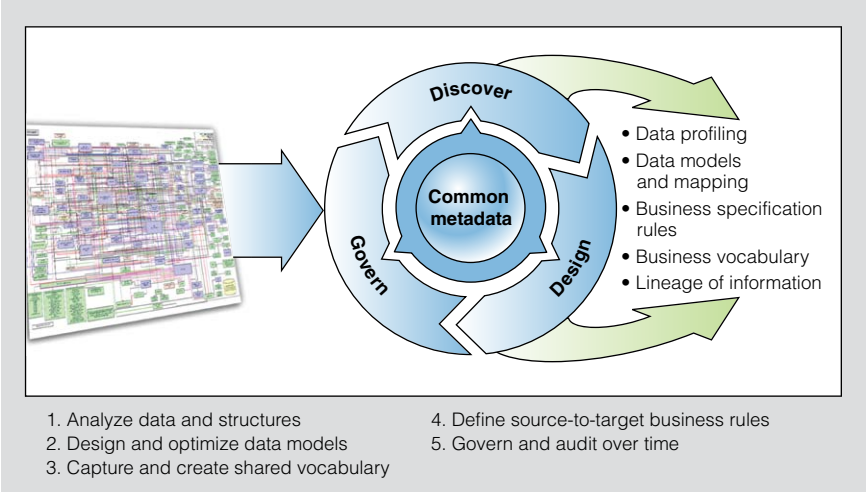
As key stakeholders, these users typically manage all of the information pertaining to various projects—including business requirements and technical specifications—in multiple, disconnected tools and processes. Managing and consolidating this information is often a manual process that is prone to errors, misinterpretations (especially when business and IT staff are replaced) and redundant efforts.

With InfoSphere Foundation Tools, users can take ownership of critical assets and monitor information quality over time, helping to build greater trust between business and IT and maximize the flow and use of information to optimize business processes. InfoSphere Foundation Tools enable these individuals to collaborate more effectively and build a common, auditable blueprint to ensure success.

Discover, design and govern trusted information with InfoSphere Foundation Tools

The process of discovering, designing and governing the IT infrastructure to create trusted information comprises five key stages (see Figure 4). InfoSphere Foundation Tools provide critical support each step of the way.

Figure 4: InfoSphere Foundation Tools support the five-stage process of turning enterprise data into trusted information.



1. ***Analyze data and structures:*** *The first step is to analyze existing enterprise data sources to understand the structure, content and quality of the information available. Existing documentation may be out-of-date and must be validated to ensure accuracy, consistency and completeness. InfoSphere Foundation Tools are designed to work in a heterogeneous IT environment comprised of IBM and non-IBM information sources.*
2. ***Design and optimize data models:*** *The second step is to build a data model for your project requirements, such as a datamart or an enterprise-wide data warehouse. You can alternatively update and optimize existing structures that are already in*

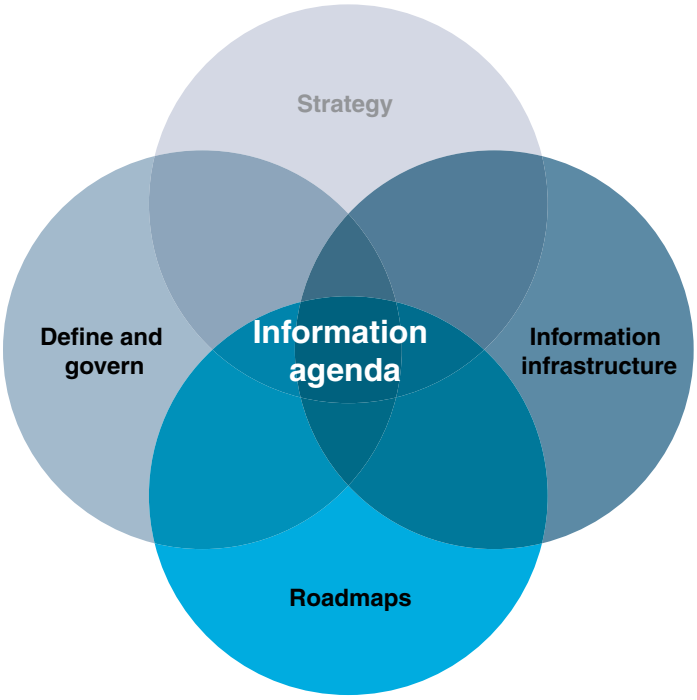
place within your organization. The IBM Industry Models may also be utilized as an add-on to InfoSphere Foundation Tools to help accelerate this process and allow you to take full advantage of years of IBM experience and industry best practices.

3. ***Capture and create shared vocabulary:*** *The third step is to capture and create a shared, enterprise-wide common vocabulary that accurately reflects your business hierarchies. Deploying a common glossary of terms for enterprise adoption is a fundamental underpinning of data governance initiatives and is vital to ensuring that business and technical users fully understand the meaning of critical enterprise assets. This step also includes classifying where these business definitions are represented by physical assets across the enterprise.*
4. ***Define source-to-target business rules:*** *The fourth step is to leverage the information learned when analyzing the enterprise data source systems, along with the new data models and business definitions, to create specific business rules. These business rules describe how to transform data sources into a consumable format for the target application and required uses.*
5. ***Govern and audit over time:*** *The fifth step is to implement repeatable processes to govern information quality over time. This includes using a standard process to define common enterprise terminology; tracking business specification documentation for historical auditing purposes; and applying approved naming conventions and data quality auditing rules to key areas. For example, IT departments might extract critical information about products, customers and suppliers from various repositories and applications on a weekly basis and assess the quality of that information using InfoSphere Foundation Tools against best-practice KPIs.*

Powering your information agenda with InfoSphere Foundation Tools

InfoSphere Foundation Tools provide an entry point to an information agenda and are one of several components that also play a critical part in implementing an information agenda. The combination of InfoSphere Foundation Tools, IBM Industry Models and tools such as extract, transform, load (ETL) and data quality products is instrumental in powering three of the four components of an information agenda: define and govern, information infrastructure and roadmaps (see Figure 5).

Figure 5: InfoSphere Foundation Tools and IBM Industry Models address three of the four components of an information agenda: define and govern; information infrastructure; and roadmaps.



"The difference between being able to evaluate millions of rows with hundreds of columns versus hundreds of rows with half the number of columns in two hours is huge. This wouldn't be possible without [InfoSphere] Information Analyzer."

–Kevin Janes
Senior Solutions Architect
Shared Health

Information infrastructure: Assess and define the information landscape

The first step in defining the information infrastructure is to inventory and understand the current information ecosystem across the enterprise. *InfoSphere Information Analyzer* profiles and analyzes source data from multiple enterprise repositories to understand the structure, quality and content of information. It also creates valuable metadata as a part of the data profiling process, identifies new data objects and enables in-depth analysis within and across heterogeneous information sources. The assessment can span the entire enterprise information ecosystem: various databases, files and enterprise applications. Without InfoSphere Information Analyzer, this process can be manual, time-consuming and error-prone.

InfoSphere Business Glossary and IBM Industry Models are instrumental to describing and shaping the target information infrastructure. InfoSphere Business Glossary establishes a consistent, common, enterprise-wide vocabulary to describe information assets in business terms. It is used to help improve collaboration between business users, IT analysts and data architects and to manage business terms and hierarchies. The IBM Industry Models are especially useful for customers who are looking to standardize data and business process models across their organization: they can be directly imported into InfoSphere Business Glossary using *Rational Data Architect* to form the basis of a common vocabulary and description of the physical data structures.

InfoSphere Business Glossary Anywhere is a desktop client designed to expose business glossary contents from within any application. It allows business analysts, architects and all end users to access business definitions from within a desktop application, including Microsoft® Office products, business intelligence products, Web browsers and others, by clicking on a term to trigger a pop-up window containing the definition. Content accessibility is an important part of creating a mature information infrastructure.

InfoSphere FastTrack builds upon the InfoSphere Information Analyzer assessment, providing business and data analysts with a method to capture and define business source-to-target specifications based on the results of data profiling. A desktop design tool, InfoSphere FastTrack creates common, familiar mapping documentation to help streamline the required collaboration between business analysts, data modelers and developers and create a solid information infrastructure with the most complete business requirements.

InfoSphere Metadata Workbench is critical to understanding the end-to-end information infrastructure built using the InfoSphere Foundation Tools. It provides an intuitive graphical user interface (GUI) for browsing information assets and understanding data lineage (where each field comes from), as well as conducting what-if analysis to predict the impact of application changes. Using InfoSphere Metadata Workbench, users can learn more about

relationships among business terms, data model entities and technical and reporting metadata—a foundation step for building and maintaining a map of your enterprise information infrastructure.

Rational Data Architect helps simplify data modeling and integration design activities by enabling architects to discover, model, visualize and relate diverse and distributed data assets. Rational Data Architect also acts as a modeling gateway to InfoSphere Foundation Tools, interchanging glossary and physical metadata. The tight integration between the IBM Industry Models, Rational Data Architect and InfoSphere Foundation Tools allows organizations to exploit industry-specific business and technical metadata to accelerate data integration projects. For example, the IBM Industry Models and Rational Data Architect physical schemas can be shared across the other InfoSphere Foundation Tools, and business or glossary definitions from the IBM Industry Models and Rational Data Architect can be used to populate InfoSphere Business Glossary to share common definitions across the enterprise.

Define and govern: Govern and manage information assets

InfoSphere Business Glossary is the starting point for designing information governance. IT and line-of-business (LOB) leadership may assign data stewards to business terms, categories or assets; using InfoSphere Business Glossary, business users can visually browse business terminology and relationships

"We're in the business of integrating healthcare data that is aggregated from numerous sources and in numerous formats, and translating that data into actionable information. To achieve this, we need scalable, robust tools that complement practice workflows and deliver immediate value."

–Hugh Hale
Vice President and Chief Technology Officer
Shared Health

and provide direct feedback to data stewards. Data stewards can then create and manage additional business terms and categories to support the end-state information architecture, while still understanding the relationship and impact to the current ecosystem.

InfoSphere Metadata Workbench helps support information governance by providing reporting and traceability on InfoSphere data movement, modeling and business intelligence applications, as required by regulations such as the Sarbanes-Oxley Act and Basel II. During the design phase of an information agenda, IT personnel can analyze and view the impact of changes to the current information model, avoiding potentially disruptive modifications to existing processes. InfoSphere Metadata Workbench relates design and operational and business metadata to drive better understanding and higher trust between IT and LOB employees.

InfoSphere FastTrack enables the centralization and tracking of all business specification requirements from inception through design and eventual fulfillment. Business mapping requirements can be re-used from the repository, and it can serve as an audit trail for design decisions made during the development process or provide historical reporting. InfoSphere FastTrack can also translate these business requirements into InfoSphere integration jobs, bridging the gap between the business analyst and the integration developer.

Roadmaps: Accelerate your path to the information agenda

IBM Industry Models can help accelerate the information agenda roadmap by providing best-practice data, business process models and enterprise business definitions that form the basis for the desired end-state. These models are a culmination of successful business transformation across hundreds of customers in multiple industries and are flexible enough to extend and accommodate the unique needs of an organization. They are designed for heterogeneous information ecosystems and support both IBM and non-IBM applications and infrastructure. The models include KPIs and reports to help track progress toward achieving the business goals of the information agenda. Migration and planning tools are also a part of the IBM Industry Models, to facilitate and speed the transition from the current state to the desired future state.

InfoSphere FastTrack provides a single location to store and manage all business specifications related to the information agenda, from initiation to deployment. It can also generate ETL jobs to transform information from its source type to the format required by the target systems. InfoSphere FastTrack essentially provides a seamless path from business requirements to integration logic and rules, helping to accelerate information integration and the information agenda roadmap.

"IBM is a definite leader in heterogeneous data integration and complied with our requirements in all areas of privacy and security."

–Naomi Rafael
Senior Database Administrator
Melbourne Health

Case in point: Melbourne Health

Today, IBM InfoSphere Foundation Tools help organizations from a broad variety of industries design, manage and govern their information architectures. For example, Melbourne Health, a public health provider in Victoria, Australia, needed to compare individual treatment plans to improve overall patient outcomes and integrate data across heterogeneous sources to drive collaboration among organizations and expansion into new research areas.

To help clinicians derive greater value from existing data, Melbourne Health led a collection of medical research organizations to launch BioGrid Australia. This award-winning program enables clinical researchers to access data from disparate databases across multiple disease types at multiple institutions. Data from more than 80,000 patients and 25 million records from up to 25 years of research are now available to authorized clinicians. The data is located in a virtual repository that can be linked with publicly available research and genetic profiling data. Researchers can correlate information from a variety of sources—including population-based health records, clinical research, disease sub-specialties, gene expression, protein expression and genotypes—to gain greater insight into the efficacy of current treatments and study how different treatments might affect different people.

"Researchers can stratify collaborative research data in a way that was impossible before. We've obtained AU$2.2 million in research grants from Australian health organizations as a result."

–Robert Merriel
Director of Business Development
Melbourne Health

IBM InfoSphere Information Server—including IBM InfoSphere DataStage®, IBM InfoSphere Federation Server, IBM InfoSphere Business Glossary and IBM Rational Data Architect—provides the core toolset to help Melbourne Health integrate information for the BioGrid project. InfoSphere DataStage enables the collection, integration and transformation of Melbourne Health's disparate data into a single research repository. InfoSphere Federation Server provides authorized researchers with virtualized access to these local research repositories, presenting them as if they were a single source.

However, it is InfoSphere Business Glossary that enables the understanding of BioGrid data for community collaboration by allowing staff to create, manage and share information about the types of research and data. Researchers can browse the industry-standard categories of data available by description and understand the technical composition of those structures. InfoSphere Business Glossary manages a significant amount of data that describes more than 25 databases, 1,000 tables and 10,000 business terms.

Two administrators are currently responsible for the glossary, in addition to the project director. The annotations are supplied by the six BioGrid data managers in consultation with the data owners. In the near future, the organization plans to train data owners to annotate their glossary terms online. Additionally, Rational Data Architect enables database administrators to quickly design and replicate database structures when needed so that researchers can begin their studies without unnecessary delays.

"Rational Data Architect helped us to semantically integrate all our critical data sources and to design a business information model within two weeks. Previously, this process would have taken more than two and a half months."

—Antoine Proult
Associated IT Architect and Senior Consultant
ACP Conseil

Why IBM InfoSphere Foundation Tools?

IBM InfoSphere Foundation Tools provide a unique combination of capabilities to help organizations define and manage their data. The tools help organizations discover and categorize new information sources; model and map data schemas; create business rules; establish and maintain data stewardship; manage business vocabulary and relationship hierarchies; and centralize this information in a shared repository to facilitate active collaboration between business and IT.

InfoSphere Foundation Tools are designed as an open toolkit that works with any IBM or non-IBM data source, business intelligence tool or data warehouse tool—or in conjunction with the tools' own comprehensive set of integration products. As a result, InfoSphere Foundation Tools can be deployed in a heterogeneous IT environment, helping you leverage existing IT investments as you create your new information architecture.

The tools can be deployed as a bundle or as stand-alone products, and IBM has industry-specific accelerators and other best practice assets to help you deploy InfoSphere Foundation Tools within your organization. IBM also provides best-practice assets for each product to help you implement your projects quickly.

IBM has the combination of mature, proven capabilities and the experience of successfully implementing the toolset for hundreds of worldwide customers across many industries: government, healthcare, energy and utilities, industrial, manufacturing, banking, insurance, financial services and retail. In addition, IBM has designed industry-specific Information Agenda Guides and workshops to help you understand where InfoSphere Foundation Tools can add value in your organization.

InfoSphere Foundation Tools and InfoSphere software portfolio: Enabling trusted information for your information agenda

InfoSphere Foundation Tools help organizations build their information agenda projects with minimal churn between business and IT teams, less risk and better results. The toolkit is also part of the larger IBM InfoSphere software portfolio, which enables organizations to maintain data quality over time, manage accessibility and structure and deliver information to the systems and decision makers that need it.

InfoSphere software provides a full set of capabilities for integrating and managing a complete and accurate view of information across your systems. The portfolio is flexible enough to allow you to start anywhere and mix and match InfoSphere software components with components from other

vendors. Anchored by InfoSphere Information Server, InfoSphere Master Data Management (MDM) Server and InfoSphere Warehouse, the InfoSphere portfolio provides a complete, integrated and easy-to-deploy platform to help address a wide range of information needs.

IBM InfoSphere Information Server

InfoSphere Information Server is a flexible data integration platform designed to deliver trusted information on demand for business initiatives. It integrates information from heterogeneous systems and helps transform data into actionable information. A parallel processing foundation allows InfoSphere Information Server to easily scale to meet the needs of even the most demanding environments, support growing data volumes and shrink processing windows with batch and real-time processing.

The unique, metadata-driven design of InfoSphere Information Server helps to align business goals with IT activities, providing a consistent understanding of business context by capturing business vocabulary and specifications and using them to automate development tasks. It also provides deeper insight into data by tracking its end-to-end lineage. These capabilities help improve overall project productivity by promoting collaboration during development and creating a set of reusable assets to drive ongoing value across multiple information projects.

IBM InfoSphere MDM Server

MDM maintains a consistent "master" record of key business entities such as customer, supplier, product, location or account and ensures that all instances of this data are synchronized across all systems at all times. InfoSphere MDM Server provides an operational data hub for all types of master data. It centralizes and synchronizes this data across heterogeneous systems through a library of more than 800 pre-packaged business services. The goal is to provide a single trusted source of information for business processes so that organizations can confidently use enterprise data to help create new revenue opportunities, reduce redundancies and enact other strategies to facilitate growth. InfoSphere MDM Server also leverages the data quality and transformation services from InfoSphere Information Server to provide consistent processing rules across multiple applications.

IBM InfoSphere Warehouse

InfoSphere Warehouse delivers a powerful and scalable foundation for your data warehouse. It includes data mining, text analytics, data archiving and compression, performance and workload management capabilities. Online analytical processing (OLAP) capabilities enable users to build logical datamarts without requiring a separate physical server and provide seamless support for IBM Cognos and other BI and reporting tools. The flexibility of logical datamarts enables InfoSphere Warehouse to dynamically react as

information changes, which helps reduce overall IT cost and drive business insight. InfoSphere Warehouse can also call directly upon transformation services from InfoSphere Information Server, providing the utmost in flexibility and performance.

IBM is focused on accelerating value with the InfoSphere software portfolio, both through the technology and through industry-specific accelerators. InfoSphere Foundation Tools provide the starting point to create a foundation for trusted data and enhance your information agenda; the InfoSphere software portfolio provides an integrated, scalable foundation to help protect your investment.

For more information

To learn more about InfoSphere Foundation Tools, the InfoSphere software portfolio and IBM Information Agenda, please contact your IBM sales representative or IBM Business Partner, or visit:

- **ibm.com**/software/data/information-agenda/foundation-tools.html
- **ibm.com**/software/data/information-agenda
- **ibm.com**/software/data/infosphere

IBM Software Group
Route 100
Somers, NY 10589

Produced in the United States of America
October 2008

TAKE BACK CONTROL WITH **Information Management**

IMW14100-USEN-00