

## IBM Information Server, Version 8.0: Understanding and analyzing your data

---

### Highlights

---

- **Create a greater understanding of data source structure, content and quality**
- **Ensure healthy data quality throughout the project life cycle**
- **Eliminate the risk and uncertainty of proliferating bad data throughout the enterprise**

### Understanding and addressing data quality: A critical early step

What price tag would you place on the value of truly understanding the data sources of your data-driven project? According to industry analysts, over 80 percent of IT projects either fail or overrun their budgets due to a lack of data understanding.

IBM® Information Server can help you quickly and easily understand your data by providing data profiling as well as monitoring capabilities. These insights can help you accelerate information-based projects—from data warehousing to infrastructure consolidation to master data management and data governance.

Data profiling (data source analysis) helps you understand your data by revealing insightful information about the structure, content and quality of your data sources. Data quality is a core discipline of IBM Information Server because it enables the delivery of consistent, reusable, trusted information. The data quality continuum includes data profiling, standardization, matching and data quality monitoring. IBM Information Server

delivers them all. For more information on standardization and matching, visit [ibm.com/software/data/integration/qualitystage](http://ibm.com/software/data/integration/qualitystage)

### IBM Information Server inspires confidence in data quality

IBM Information Server automates data profiling, analysis and monitoring, helping to not only simplify information integration, but also inspire greater confidence in the quality of information delivered. IBM Information Server allows you to meet the demanding requirements of your enterprise today while achieving the objectives that will help sustain your growth for tomorrow. By using the data profiling capabilities of IBM Information Server in the early phases of your data integration projects, you can:

- *Expedite delivery of data-driven projects*
- *Minimize costs and resources of critical data integration projects*
- *Eliminate the risk and impact of proliferating incorrect and inaccurate data*
- *Help ensure the timely delivery of trusted information*

IBM Information Server profiles data by performing complete source-system analysis, including column and table analysis, primary and foreign key analysis, relationship analysis and redundancy analysis on your heterogeneous data sources. It also features extensive reporting capabilities, robust security, scheduling and a customizable dashboard, all in a state-of-the-art intuitive GUI (see Figure 1).

IBM Information Server data profiling and analysis features include:

### Comprehensive data analysis—

A complete set of metrics offers a complete picture of your data from every possible angle and enables users to immediately document all discovered data anomalies including data type, minimum, maximum, average, count and nulls, and others.

**Drill down—**View individual records from profiling results in real time. For example, if you discover an invalid value in a column, you can easily drill down to the actual record for further investigation.

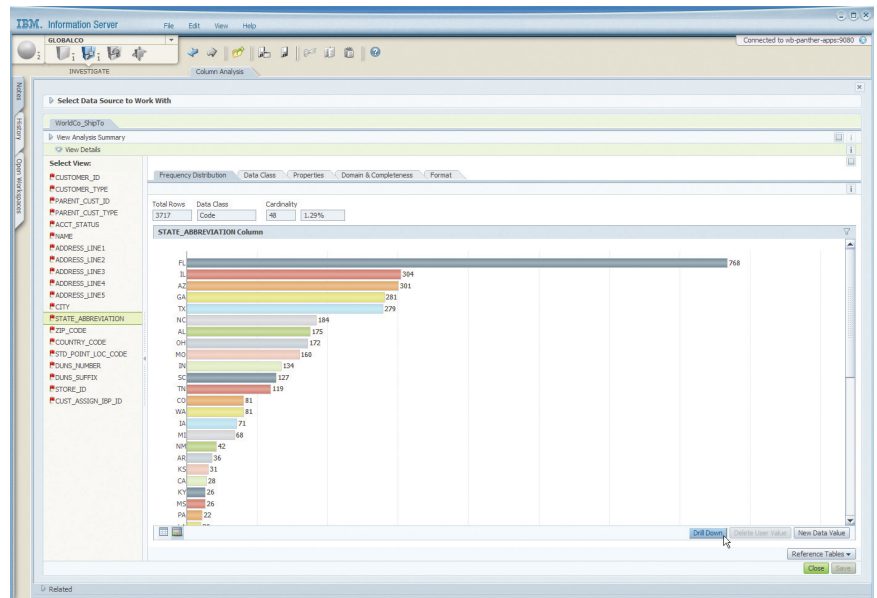


Figure 1. Frequency distribution results from column analysis

**Extensive reporting—**Over 35 reports are ready to use out of the box, customizable for your environment. Visual representation of analysis, trends and metrics and incorporation of user notes all help you understand results quickly and efficiently.

**Intuitive user interface—**A rich, visual and graphically enabled user interface with intuitive navigation experience promotes the understanding of information and improves usability and productivity. A task-based approach to navigation helps you quickly perform

what you need to do, when you need to do it (see Figure 2).

**Native parallel data profiling—**Profile high volumes of data easily without any added configuration.

**Secured analysis—**Project-, role- and user-based security controls access to sensitive analytical information. Controls can be as open or as restrictive as needed to meet compliance requirements.

**With IBM WebSphere® Information Analyzer, leverage an understanding of source systems to uncover anomalies including:**

- Missing data
- Inaccurate and inconsistent data
- Duplicate data

**Collaborative multiuser environment**—With support for multiple analytical reviews, as well as asynchronous profiling, a project-based environment allows more than one user to work in a project-based context.

**Baseline analysis**—Compare results from current profiling sessions with the results from a prior profiling session, gaining insight into what may have changed in structure or content over time.

**Integrated metadata**—Share analytical results and profiling notes across all IBM Information Server modules.

**Broad support of data sources**—

Profile IBM DB2®, Oracle, Microsoft SQL Server, Sybase, IBM Informix®, Microsoft® Access, text files, Open Database Connectivity (ODBC) data sources and more.

**Data profiling within a unified platform**

Leveraging a Service Oriented Architecture (SOA) to unlock information from individual silos, WebSphere Information Analyzer, a product module of IBM Information Server, delivers data profiling and analysis functions within the context of a complete information integration platform,

enabling more accessibility and consistency throughout the enterprise.

Active metadata across IBM Information Server simplifies the collection and management of metadata across the entire integration spectrum. Within the WebSphere Information Analyzer module, profiling results are stored in the common metadata repository. The result? Profiling outcomes pertinent to IBM WebSphere DataStage® and IBM WebSphere QualityStage™ users are instantly accessible without the need to export/import the data.

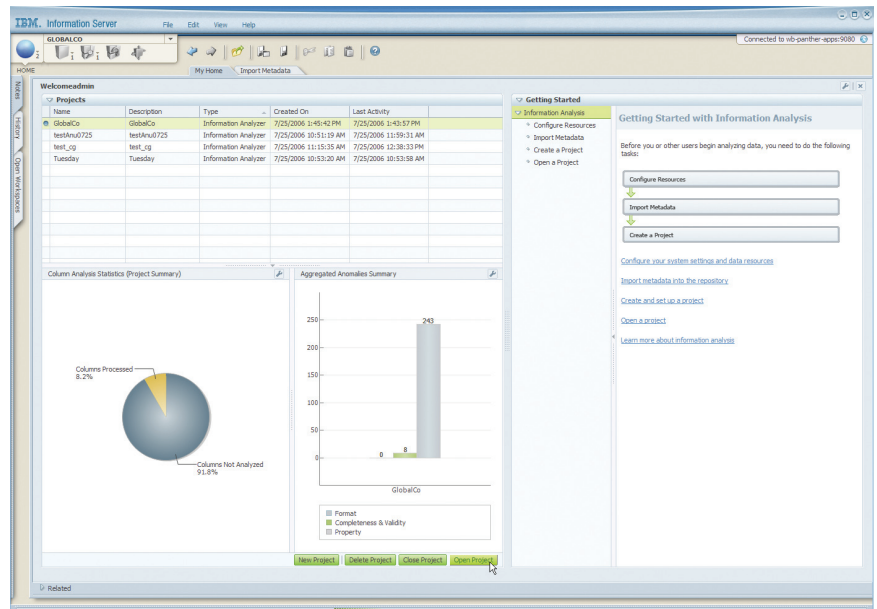


Figure 2. Rich, task-based graphical user interface with customizable dashboard



### **IBM Information Server delivers trusted information**

Organizations face an information challenge beginning with locating information, getting it when it is needed in the form needed, and once it is found, discerning further insights from it. Information validity and control are additional concerns. The challenges only mount if businesses cannot ensure access to authoritative, consistent, timely and complete information.

IBM Information Server is a revolutionary new software platform that helps you derive more value from the complex, heterogeneous information spread across your systems. It enables your organization to integrate disparate data and deliver trusted information wherever and whenever needed, in line and in context, to specific people, applications and processes. It helps business and IT personnel collaborate to understand the meaning, structure and content of any type of information across any source. It provides breakthrough productivity and performance for

cleansing, transforming and moving this information consistently and securely throughout the enterprise, so it can be accessed and used in new ways to drive innovation, increase operational efficiency and help lower risk.

### **System requirements**

WebSphere Information Analyzer—the IBM Information Server module that delivers information profiling and analysis functions—supports the following operating systems: IBM AIX®, Red Hat Enterprise Linux®, SuSE Enterprise Linux, Solaris, HP-UX and Microsoft® Windows®. For current, detailed hardware and software system requirements for IBM information integration products, visit [ibm.com/software/data/integration](http://ibm.com/software/data/integration)

### **For more information**

To learn more about IBM Information Integration Solutions, contact your IBM marketing representative or IBM Business Partner, or visit [ibm.com/software/data/integration](http://ibm.com/software/data/integration)

© Copyright IBM Corporation 2006

IBM Software Group  
Route 100  
Somers, NY 10589

Printed in the United States of America  
November 2006  
All Rights Reserved

AIX, DataStage, DB2, IBM, the IBM logo, Informix and WebSphere are trademarks of International Business Machines Corporation in the United States, other countries or both.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft and Windows are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product or service names may be trademarks or service marks of others.

References in this publication to IBM products or services do not imply that IBM intends to make them available in all countries in which IBM operates. Offerings are subject to change, extension or withdrawal without notice.

All statements regarding IBM future direction or intent are subject to change or withdrawal without notice and represent goals and objectives only.

The information contained in this document is provided for informational purposes only. While efforts were made to verify the completeness and accuracy of the information contained in this document, it is provided "as is" without warranty of any kind, express or implied. In addition, this information is based on IBM's current product plans and strategy, which are subject to change by IBM without notice. IBM shall not be responsible for any damages arising out of the use of, or otherwise related to, this document or any other documents. Nothing contained in this document is intended to, nor shall have the effect of, creating any warranties or representations from IBM Software.

**TAKE BACK CONTROL WITH** **Information Management**