

ORACLE DATA PROFILING

ORACLE DATA PROFILING

KEY FEATURES

- Automated entity reverse-engineering and analysis
- Comprehensive statistics and advanced profiling techniques
- User interface provides a natural drill-down approach
- Keys, functional dependencies, and joins discovery and analysis
- User-defined rules
- Built-in documentation and annotation system
- Report generation and export for third-party tools
- The ability to monitor quality evolution over time
- E-mail alerts to key users about quality issues
- Integration with Oracle Data Quality

Most business software systems contain inaccurate data, information inconsistencies, and other anomalies. Oracle Data Profiling includes data investigation and quality assessment tools to help you find these issues. It allows business users to get a clear picture of their data quality, and monitor, track, and govern information over time.

Oracle Data Profiling

Every enterprise software system contains inaccuracies, inconsistencies, and anomalies. Internal workarounds and assumptions usually compensate for these errors, but consolidating and integrating data frequently bring these data quality issues to light. When used for a master data management, data warehousing, or other data integration project, these data sources can become chaotic and problematic due to conflicting data assumptions. Further complicating the matter, regulatory compliance initiatives (such Sarbanes-Oxley, U.S. Patriot Act, and Basel II) require tracing the source of the data used in financial reports, as well as examination, tracking (through snapshots), and certification of the state and quality of the business data.

Oracle Data Profiling is a data investigation and quality-monitoring tool. It allows business users to assess the quality of their data through metrics, discover or infer rules based on this data, and monitor the evolution of data quality over time.

Powerful Data Analysis

Oracle Data Profiling provides a set of powerful features for investigating data quality, including

- **Entity Discovery and Analysis**—Oracle Data Profiling collects metadata and data from the sources and analyzes it to render comprehensive information and statistics, such as attribute lengths, maximum and minimum values, value distributions, patterns, data types, and so on. It automatically applies advanced profiling techniques to identify potential problems with data fields such as nonconformed zip, customer, or product codes; misspellings; duplications; and punctuation issues.
- **Natural Drill-Down**—The user interface provides a natural drill-down approach to browse through the analysis results and graphical views.
- **Keys, Functional Dependencies, and Joins Discovery and Analysis**—Oracle Data Profiling detects and presents potential keys, along with their degree of uniqueness, while identifying duplicates and other inconsistencies. It also detects functional dependencies between attributes within a given entity (a shipped order should have an invoice number, for example), as well as relationships between entities (joins). Quality analysts can also create these

ORACLE DATA PROFILING

types of quality rules within the user interface.

KEY BENEFITS

- Powerful data analysis enables automatic detection of data quality issues and deduction of new quality rules
- Intuitive user interface for business users
- The Time Series feature allows for data quality monitoring over time and e-mail alerts enable emergency cases handling



Figure1. Oracle Data Profiling attribute drill-down and join Venn diagram

RELATED PRODUCTS AND SERVICES

- Oracle Data Integrator
- Oracle Data Quality for Data Integrator

Reporting, Monitoring, and Emergency Handling

Oracle Data Profiling allows data quality specialists to document and annotate their findings, define business rules, and create data quality reports or metadata and data exports for third-party tools.

The Time Series feature of Oracle Data Profiling allows users to evaluate data quality over time by performing regular assessments of the data quality. The state of the data is preserved with all analysis findings.

Oracle Data Profiling also uses e-mail notifications to warn business users when certain service-level requirements are not met (such as when threshold values are exceeded, business rules are violated, and so on). Data quality stewards can then review and evaluate these errors in more detail.

Integration in the Data Quality Process

Oracle Data Profiling is tightly integrated with Oracle Data Quality and shares the same user interface. Once the data investigation is complete and the data quality challenges are correctly identified, business and IT users can automatically generate their data quality project for use with downstream extract-load-transform (ELT) and service-oriented architecture (SOA) processes.

Copyright © 2007, Oracle Corporation and/or its affiliates.

All rights reserved. Published in the U.S.A. This document is provided for information purposes only, and the contents hereof are subject to change without notice. This document is not warranted to be error-free, nor is it subject to any other warranties or conditions, whether expressed orally or implied in law, including implied warranties and conditions of merchantability or fitness for a particular purpose. We specifically disclaim any liability with respect to this document, and no contractual obligations are formed either directly or indirectly by this document. This document may not be reproduced or transmitted in any form or by any means, electronic or mechanical, for any purpose, without our prior written permission.

Copyright © 2007, Oracle is a registered trademark of Oracle Corporation and/or its affiliates. Other names may be trademarks of their respective owners.