# ORACLE ENTERPRISE DATA QUALITY PRODUCT FAMILY

**ORACLE**®

**FUSION MIDDLEWARE**

DATA QUALITY

The Oracle Enterprise Data Quality family of products helps organizations achieve maximum value from their business-critical applications by delivering fit-for-purpose data. These products also enable individuals and collaborative teams to quickly and easily identify and resolve any problems in underlying data. With Oracle Enterprise Data Quality products, customers can identify new opportunities, improve operational efficiency, and more efficiently comply with industry or governmental regulation.

## KEY FEATURES

**Profile and Audit**

• Profiling uncovers and quantifies hidden data problems

• Audit rules measure the quality of data against your business rules

• Easy-to-use interface built for business- IT collaboration

**Parsing and Standardization**

• Transform and standardize data such as names, addresses, dates and phone numbers

• Extract structured information from freeform text

• Prepare data to optimize its value to business applications

**Match and Merge**

• Match parties at the individual, group or household level

• Works for individuals and corporate entities

• Fully flexible rules easily tuned to suit your business using pre-build templates

## Purpose-Built Data Quality

Ever since there have been databases and applications, there have been data quality problems. Unfortunately all those problems are not created equal and neither are the solutions that address them. Some of the largest differences are driven by the data type, or domain, of the data in question. The most common data domains in data quality are customer (or more generally, party data including suppliers, employees, etc.) and product data. Oracle Enterprise Data Quality products recognize these differences and provide purpose-built capabilities to address each.

Quick to deploy and easy to use, Oracle Enterprise Data Quality products bring the ability to enhance the quality of data to all stakeholders in any data management initiative. Oracle Enterprise Data Quality products cover:

• Profiling, Audit and Dashboards

• Parsing and Standardization

• Match and Merge

• Case Management

• Address Verification

• Product Data Capabilities

Each of these capabilities is described in the following sections.

## Profile, Audit and Dashboards

Profiling current data provides a basis for understanding data quality issues and a foundation for building data quality rules for defect remediation and prevention. It provides the ability to understand your data, highlighting key areas of data discrepancy,

**ORACLE**®

**Address Verification**

• Comprehensive global coverage for address verification and geocoding

• Adds geocode to city or postal code for over 240 countries

• Robust parsing, transliteration and standardization for any address, anywhere

**Product Data Extension**

• Semantic-based recognition with auto- learning

• Ability to handle multiple product categories and extreme variability

• Item classification, attribute extraction, and standardization

• Standardize descriptions in any language

• Exact, similar, and related matches

• Ability to merge records based on survivorship rules

helping you to analyze the business impact of these problems, learn from historical analysis, and define business rules directly from the data. This avoids preconceptions of how the data fields relate to each other and quickly identifies weaknesses in existing business processes and technology implementations.

Oracle Enterprise Data Quality enables business teams to profile large volumes of data from databases, spreadsheets, and flat files with ease. Phrase profiling—Oracle's unique approach to understanding text data—helps you to identify key information buried in free-format text data fields. Profiling gathers intelligence and statistics about your data but does not change it. Systematic audit reviews detect key quality metrics, missing data, incorrect values, duplicate records and inconsistencies.

Results of these profiling and audit processes are presented in easy-to-understand executive dashboards. Using a web browser, workers and managers can monitor and review ongoing data quality against defined metrics. Data quality dashboards allow problems to be quickly identified and dealt with before they start to cause significant business impact. Graphical views show data quality trends over time, helping your organization protect its investment in data quality, by giving visibility to the right people.

Profiling, audit and dashboards provide the core understanding and visibility required to start a tactical data quality project, or an enterprise-wide data governance initiative.

## Parsing and Standardization

Oracle Enterprise Data Quality provides a rich palette of functions to transform and standardize data using easily managed reference data and simple graphical configuration. In addition to functions for basic numeric, string, and date fields, functions for contextual data such as names addresses and phone numbers are provided. Users can also quickly configure, package, share and deploy new functions that encapsulate rules specific to their data and industry without any coding.

Text data is very rarely available in a completely neat and ordered fashion. Typical problems include:

• Constructed fields, where a customer ID may be made up of a location code, a customer reference, and an account manager code
• Misfielded data, such as names, comments, or telephone numbers in address blocks
• Poorly structured data such as addresses, where data can flow from one field to the next.
• Notes fields that store information that the data structure doesn't support, but that contain useful semi-structured data that normally cannot be analyzed or extracted

All of these problems can be solved using Oracle Enterprise Data Quality. Using a data-driven approach to rapidly tag or describe data, it can manipulate a single record by it parsing into multiple structured elements (and, if required, records) and standardize results according to predefined rules. Innovative parsing and phrase analysis technology uniquely allows you to find hidden knowledge within any text field and create rules to standardize it into structured data.

Oracle Enterprise Data Quality can also be used to audit against defined business rules and transform data on the fly against those rules, providing a flexible and adaptable

data quality firewall.  In addition, it allows the entire assembled data quality process to be called as a real- time Web service. Results of the parsing and standardization processes can be viewed in graphical dashboards that provide a complete, accurate, and accessible view of your business world.

## Match and Merge

Matching is a key component of many data quality projects, and can be used to support different activities such as de-duplication, duplicate prevention, consolidation, customer data integration (CDI) and master data management (MDM). Oracle Enterprise Data Quality provides powerful matching capabilities that allow you to identify matching records and optionally link or merge matched records based on survivorship rules. Many flexible yet intuitive match processors and related capabilities are included to handle any data situation. Importantly, all matching rules are completely open and transparent, allowing a business analyst to quickly determine why a match rule may have triggered and allow them to quickly tune the match behavior as required.  Matches may be determined solely by the system, or it can be flagged as a potential match and put in a queue for manual review.

Oracle Enterprise Data Quality also includes a connector that enables easy access to data in Oracle's Siebel CRM. Audit capabilities allow you to run data quality rules and flow-control within your data quality processes. Dashboard functionality presents results of the audit processes in graphical format, while real-time Web service functionality enables whole assembled data quality process to be called as a real-time service.

## Case Management

Oracle Enterprise Data Quality will automate a large proportion of the required 'cleanup' tasks, but some data rules and validation may still not be met where there is missing or incorrect data.  In this situation, records can be placed in a 'case management' queue for subsequent manual review and remediation.  Of course, not all data errors are worth correcting, so the queue can be prioritized based on business rules, so that only the most important data is handled this way.  This Case Management facility is an ideal complement to the more 'pure' automation and is especially useful for pre-validating and remediating data as part of a system migration or consolidation where faulty data may not otherwise be detected until after go-live, thus undermining the effectiveness of the new system from day one.

## Address Verification

Many data quality problems involve names and addresses.  It is one thing to ensure an address is valid in format, but quite another to verify that it actually exists and is a real, deliverable address.  Oracle Enterprise Data Quality fills this gap and uses reference information from various postal authorities around the world to verify that addresses are 'real'.  In addition, for verified addresses, it can also return a geo-code for mapping applications.  The system can parse, standardize, transliterate and process addresses from over 240 countries – basically all the populated territories on the planet – and can handle them in structured or unstructured format, and in any character set.

## Product Data Extension

In the world or data quality, product data provides some specific challenges. The rules governing product data are specific to the category of product being described. For example, the data quality rules for resistors are different from capacitors, which are also different from switches, fasteners, and any other product category. Each product category will have different vocabulary, terms, abbreviations, valid values, and standardizations. In addition, product information is typically communicated through nonstandard description fields that must be recognized and parsed. Compounding this problem, most data quality scenarios involving product data do not cover just one category, but hundreds or thousands of product categories.

To handle this level of variability Oracle Enterprise Data Quality can use specialized semantic recognition to quickly recognize the product category and apply the correct rules based on context. Based on the context, it can also make inferences about the meaning of a particular word or phrase and 'learn' new rules and context as it goes along. Once properly recognized, product information can be transformed and standardized including classification(s), attributes, and descriptions which can be generated in any language for consumption in downstream systems.

Product data also presents specific challenges for matching and merging product records and in this case the Product Data Extension can be used to identify and extract key product information and create a standardized item record. It can additionally identify exact, similar, and related records and optionally merge them based on defined survivorship rules.

These facilities can operate in any language and includes a connector to Oracle Product Data Hub, allowing clean, standardized, de-duplicated product information to be loaded to the MDM hub.

## Integration with Oracle Master Data Management

Data quality and MDM are highly interrelated. MDM hubs need to be loaded with high quality, complete, and standardized information. Once a piece of data is transformed to conform with relevant data quality rules, that piece of information should be stored as reference data in a hub. Oracle Enterprise Data Quality products can be used with any MDM solution, but comes pre-integrated with Oracle Customer Hub and Oracle Product Hub.

## Integration with Oracle Data Integrator

When data is being moved between systems there is a distinct need to assure the quality, consistency, and overall usability of the data in question. Conversely, one could question the value of integrating data that is of unknown quality and consistency.

Oracle Enterprise Data Quality can be used with any data integration or ETL system, but is pre-integrated with Oracle's flagship product for data movement and transformation—Oracle Data Integrator. This integration enables customers to take advantage of Oracle Enterprise Data Quality products in their integration projects for simple and rapid deployment as part of a complete data integration solution.

**ORACLE®**

**Hardware and Software, Engineered to Work Together**

Oracle is committed to developing practices and products that help protect the environment