

Harte-Hanks

Informatica

Business Obj

Zoomix

Human Inference

# Data Quality Management Software Product Directory

2009 EDITION

Data Mentor

Innovative Systems, Inc.

Stalworth Inc.

Datanomic

Oracle

Dataactics

*This product brought to you by:*



SearchDataManagement.com

[INTRO](#)

[DATA QUALITY  
MANAGEMENT  
AND CHOOSING  
DATA QUALITY  
TOOLS](#)

[INDEX](#)

[BUSINESS  
OBJECTS  
\(NOW SAP\)](#)

[DATACTICS](#)

[DATAFLUX](#)

[DATALEVER](#)

[DATA MENTORS  
INC.](#)

[DATANOMIC](#)

[EPRENTISE](#)

[FUZZY!  
INFORMATIC  
\(NOW SAP\)](#)

[HARTE-HANKS  
TRILLIUM  
SOFTWARE](#)

[HUMAN  
INFERENCE](#)

[IBM](#)

[INFORMATICA](#)

[INNOVATIVE  
SYSTEMS INC.](#)

[NETRICS](#)

[ORACLE CORP](#)

[PERVASIVE  
SOFTWARE](#)

[PITNEY BOWES  
GROUP 1](#)

[STALWORTH INC.](#)

[TALEND](#)

[UNISERV](#)

[ZOOMIX \(NOW  
MICROSOFT\)](#)

[ABOUT  
SEARCHDATA-  
MANAGEMENT  
.COM](#)

[METHODOLOGY](#)

# Welcome!

Welcome to SearchDataManagement.com's Data Quality Management Software Product Directory. This directory is designed to be a valuable resource for those getting started with research or evaluating vendors in the data quality market.

Inside, you'll find basic information about the major vendors in the data quality market and the products they sell. Each listing is accompanied by a short description and a long description including limited information about functionality and product use. You'll find products for businesses of all sizes as well as products that can be deployed on-demand and on-premise. Use this list to get started with the evaluation process. For more information about any of the products or to speak to a sales representative, please visit the vendor website or product website.

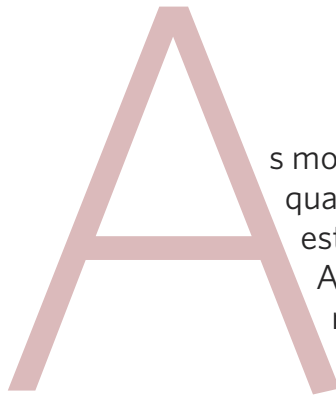
[SearchDataManagement.com](#) will launch a series of directories throughout the year to address unique segments of the data management market. Want to see your product listed in one of our directories? Go here to [submit a product](#). Need to update product or pricing information? [Email us here](#). For questions for the editors or to make suggestions for improving the directory, write to us at [editor@searchdatamanagement.com](mailto:editor@searchdatamanagement.com).

Happy shopping!



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

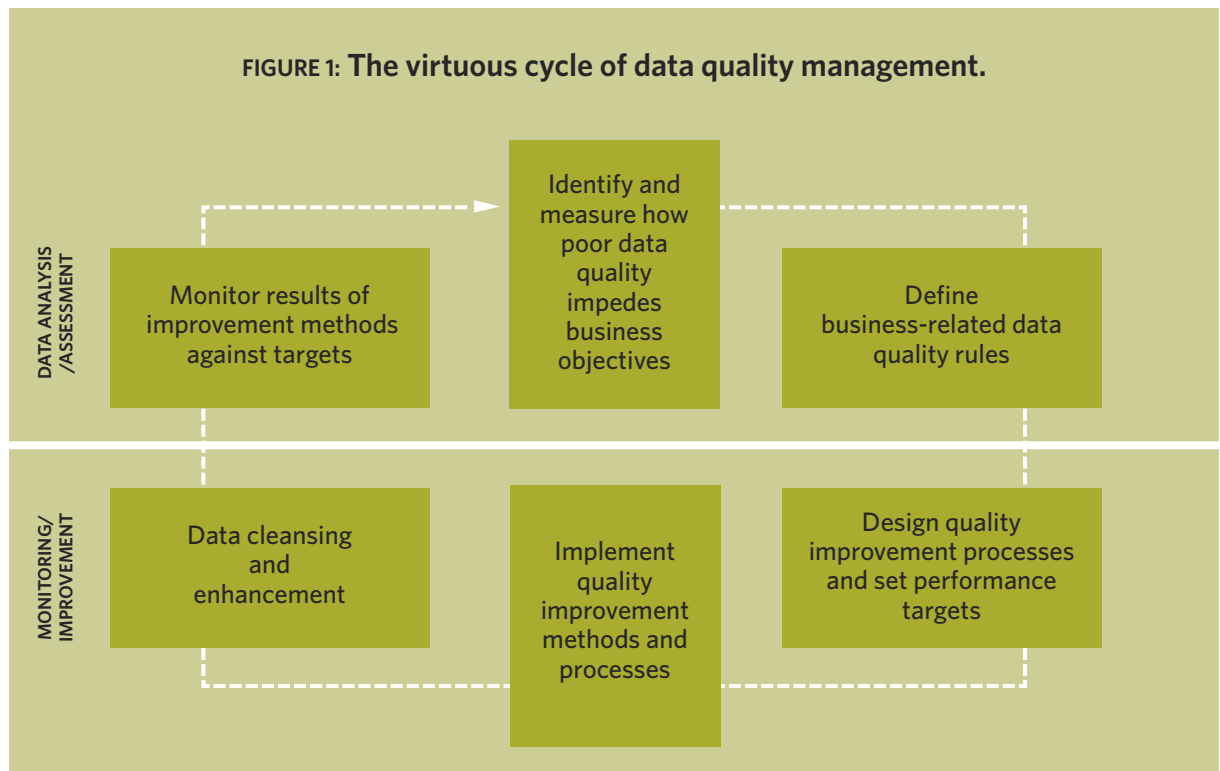
# Data quality management and choosing data quality tools: A primer



As more business analysts recognize the relationship between high-quality data and the success of the business, there is a growing interest in integrating data quality management within the organization. And while the lion’s share of the effort involves putting good data management practices in place and establishing data governance, there will always be a requirement for technology to support data quality maturity.

That being said, until relatively recently, most people equated the phrases “data quality” and “data cleansing” with the expectation that data quality tools were intended only to help identify data errors and then correct those errors. In reality, there are many techniques applied within the context of a data quality management program, with different types of tools used to support those techniques.

Data quality management incorporates a “virtuous cycle,” shown in **FIGURE 1**, essentially consisting of two phases: analysis and assessment, followed by monitoring and improvement.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

Data quality tools and technology are necessary to support both the analysis and assessment phases. Data profiling tools are used for data analysis and identification of potential anomalies, while parsing and standardization tools are employed for recognizing errors, normalizing representations and values, and some degree of data correction. These tools can be used to define data quality rules that assert validity of data and are used to flag non-conformant values and to aid the correction process.

A commonly used technology for customer and business name correction is identity resolution, which helps in linking and resolving variant representations of the same entities. After normalization and identity resolution have been performed, data enhancements such as address standardization and enhancement and geo-coding are applied. Lastly, the data quality rules can be integrated within a data quality auditing tool that measures compliance with defined data quality expectations. The results of these measurements can be fed into a data quality metrics scorecard, and if the metrics are defined in relation to the business impacts incurred by violating the expectations, that scorecard will provide an accurate gauge of how improving data quality goes straight to the bottom line.

**“Data quality tools and technology are necessary to support both the analysis and assessment phases.”**

**—DAVID LOSHIN**

### DATA PROFILING

The initial attempts to evaluate data quality are processes of analysis and discovery, and this analysis is predicated on an objective review of the actual data values. The values populating the data sets under review are assessed through quantitative measures and analyst review. While a data analyst may not necessarily be able to pinpoint all instances of flawed data, the ability to document situations where there may be anomalies provides a means to communicate these instances with subject-matter experts whose business knowledge can confirm the existence of data problems.

Data profiling is a set of algorithms for statistical analysis and assessment of the quality of data values within a data set, as well as exploring relationships that exist between value collections both within and across data sets. For each column in a table, a data profiling tool will provide a frequency distribution of the different values, providing insight into the type and use of each column. Cross-column analysis can expose embedded value dependencies, while inter-table analysis explores overlapping value sets that may represent foreign key relationships between entities, and it is in this way that profiling can be used for anomaly analysis and assessment.

The data profiling process often sheds light on business rules inherent to each business process’s use of the data. These rules can be documented and used during the auditing and monitoring activity to measure validity of data.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## DATA PARSING AND STANDARDIZATION

In any data set, slight variations in representation of data values easily lead to situations of confusion or ambiguity for both individuals and other applications. For example, consider these different character strings:

- ▶ 301-754-6350
- ▶ (301) 754-6350
- ▶ 301.754.6350
- ▶ 866-BIZRULE

All of these formats use digits, some have alphabetic characters, and all use different special characters for separation, but to the human eye they are all recognized as reasonable telephone number formats. To determine whether these numbers are accurate or to investigate whether duplicate telephone numbers exist, the values must be parsed into their component segments (area code, exchange and line) and then transformed into a standard format.

When analysts are able to describe the different component and format patterns used to represent a data object (person's name, product description, etc.), data quality tools can parse data values that conform to any of those patterns and even transform them into a single, standardized form that feeds the assessment, matching and cleansing processes. Pattern-based parsing can automate the recognition and subsequent standardization of meaningful value components.

In general, parsing uses defined patterns managed within a rules engine used to distinguish between valid and invalid data values. When patterns are recognized, other rules and actions can be triggered, either to standardize the representation (presuming a valid representation) or to correct the values (should known errors be identified).

## IDENTITY RESOLUTION: SIMILARITY, LINKAGE AND MATCHING

A common requirement for data quality management involves two sides of the same coin: when multiple data instances actually refer to the same real-world entity, as opposed to the perception that a record does not exist for a real-world entity when in fact it really does. Both of these problems indicate the need for techniques to help identify approximate matches to determine similarity between different records. In the first situation, similar, yet slightly variant representations in data values may have been inadvertently introduced into the system, while in the second situation, a slight variation in representation prevents the identification of an exact match of the existing record in the data set.

Both of these issues are addressed through a process called identity resolution, in which the degree of similarity between any two records is scored, most often based on weighted approximate matching between a set of attribute values between the two records. If the score is above a specific threshold, the two records are deemed to

- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

be a match and are presented to the end client as most likely to represent the same entity. Identity resolution is used to recognize when only slight variations suggest that different records are connected and where values may be cleansed.

Attempting to compare each record against all the others to provide a similarity score is not only ambitious but also time-consuming and computationally intensive. Most data quality tool suites use advanced algorithms for blocking records that are most likely to contain matches into smaller sets, whereupon different approaches are taken to measure similarity. In addition, there are different approaches to matching—a deterministic approach relies on a broad knowledge base for matching, while probabilistic approaches employ statistical techniques to contribute to the similarity scoring process. Identifying similar records within the same data set probably means that the records are most likely duplicated and may be subjected to cleansing and/or elimination. Identifying similar records in different sets may indicate a link across the data sets, which helps facilitate cleansing, knowledge discovery, reverse engineering and master data aggregation.

**“Consider whether or not a specific tool offering necessitates purchasing a suite of products.”**

—DAVID LOSHIN

### DATA CLEANSING AND ENHANCEMENT

Data cleansing incorporates techniques such as data imputation, address correction, elimination of extraneous data, and duplicate elimination, as well as pattern-based transformations. Data cleansing complements (and relies on) parsing and standardization as well as identity resolution and record linkage. Data enhancement is a data improvement process that relies on record linkage, along with value-added improvement from third-party data sets (such as address correction, geo-demographic/psychographic imports, list appends). This is often performed by partnering with data providers, using their aggregated data as a “source of truth” against which records are matched and then enhanced.

### DATA AUDITING AND MONITORING

The same types of data quality rules exposed through conversations with subject-matter experts and profiling can be used to describe end-user data quality expectations. Monitoring defined data quality rules and auditing the results provides a proactive assessment of compliance with expectations, and the results of these audits can feed data quality metrics populating management dashboards and scorecards.

Data profiling tools, as well as standalone auditing utilities, often provide capabilities for proactively validating data against a set of defined (or discovered) business rules. In this way, the analysts can distinguish those records that conform to defined data quality expectations from those that don’t, which in turn can contribute to baseline measurements and ongoing auditing for data quality reporting.

- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## WHAT TO LOOK FOR IN DATA QUALITY TOOLS AND VENDORS

There are two interesting notions to keep in mind when considering data quality tools. First, every organization's needs are different, depending on the type of company, industry and business processes and their corresponding dependence on the use of high-quality data. Second, while the needs may be different, the ways that those can be addressed are often very similar, although different vendors may address those issues with greater degrees of accuracy and precision (and, naturally, cost). Weighing both of these notions together, the conclusion is that what will distinguish the suitability of one product over the others is more than just functionality, especially as data quality technology becomes more of a commodity capability.

Along with functionality, consider cost, installed base, vendor stability, training and support capabilities, as well as the pool of talent that can be tapped to help integrate the tools within a governed data quality program. In addition, because there have been many corporate acquisitions within the data quality market, consider whether or not a specific tool offering necessitates purchasing a full-blown suite of products. Alternatively, one must consider the level of comfort of purchasing one component of a vendor's tools suite with the expectation that it will integrate well with other tools already in use within the environment.

## BUSINESS NEEDS ASSESSMENT AND DATA QUALITY TOOL REQUIREMENTS

The desire to acquire data quality tools should be tempered by the assessment process—too often, the technology is purchased long before the specific business needs have been determined. Therefore, it is worthwhile to perform a high-level data quality assessment with these specific objectives:

- ▶ Identify business processes that are affected by data quality issues.
- ▶ Identify the data elements that are critical to the successful execution of those business processes.
- ▶ Evaluate the types of errors and data flaws that can occur.
- ▶ Quantify business impacts associated with each of those errors.
- ▶ Prioritize issues based on potential business impacts.
- ▶ Consider the data quality improvements that can be applied to alleviate the business impacts.

While this process is presented simply above, there are many subtleties that may require additional expertise. To expedite this assessment, you may consider partnering with expert consultants that can perform the rapid assessment while simultaneously training your staff to replicate the process on other data sets.

### General requirements for data quality tools

As a result of this process, companies should arrive at a prioritized list of improvements, which should frame the discussion of requirements analysis, both from the data quality standpoint and from the systemic and environmental aspects. For example, determining that duplicated customer records lead to business risks would sug-



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

gest that duplicate elimination is advisable. This requires tools for determining when there are duplicates (identity resolution) and for cleansing (parsing and standardization, enhancement).

There are also degrees of precision, however. If your company is a mail-order vendor sending out duplicate catalogs, the risk is increased costs and lowered response, but 100% de-duplication may not be a requirement. But attempting to identify terrorists at the airport security gate may pose a significant risk in terms of passenger safety, necessitating much greater precision in identity resolution. Increased precision is likely to correspond to increased costs, and this is another consideration.

In terms of environmental and systemic aspects, one must consider how well different products can integrate within the organization’s system architecture. Hard requirements such as platform compliance are relatively easy to specify. Architectural expectations are too, such as the different deployment options such as whether the tools support real-time operations, whether they only execute in batch, or can they be integrated “in-line” are also relevant questions. In addition, as more organizations migrate toward services-oriented architectures (SOAs), determining whether the tools support services also becomes a requirement. From the business side, one must consider the license, support, training, and ongoing maintenance costs, as well as the internal staffing requirements to manage and use the products.

Because many vendors provide tools that may (or may not) address the organization’s issues, it is worthwhile to carefully delineate your business needs and technology requirements within a request for proposal (RFP). Providing an RFP provides two clear benefits. First, it clarifies your needs to the vendors so they can more effectively determine if their product will meet them. Second, it provides a framework for comparing vendor tools, scoring their relative suitability and narrowing the field. It is a good idea to include specific metrics associated with the quality of the data that can be used to compare and measure effectiveness of the products.

### **NARROWING DATA QUALITY VENDORS**

Reviewing the RFP responses will help filter out those vendors that make the grade from those whose products are not entirely appropriate to address the business needs. But to narrow the remaining vendors to a short list, set up meetings for the vendors to present their technology along with a proposal for how their products will be used to address the business needs. Again, it may be worthwhile to engage individuals with experience in data quality tools and techniques to clarify the distinctions among the vendor products, translate any “tech-talk” into terms that are understood, and to ask the tough questions to ensure that the vendors are properly representing what their products can and cannot do.

By this time, your team should be able to whittle down the field to at most three competitors. The final test is to try out the tools yourself—arrange for the installation of an evaluation version of the product and run it over your own data sets. Having specified a benchmark data set for comparison, one can compare not just how well the products perform but also the ease of use and adoption by internal staff.



## SPECIFIC REQUIREMENTS

The full details of what can be expected from the data quality tools described is beyond the scope of this article, but this table provides some high-level, “no-miss” capabilities for each of the tools described.

TECHNOLOGY	CORE CAPABILITIES
Data profiling	<ul style="list-style-type: none"> <li>Column value frequency analysis and related statistics (number of distinct values, null counts, maximum, minimum, mean, standard deviation)</li> <li>Table structure analysis</li> <li>Cross-table redundancy analysis</li> <li>Data mapping analysis</li> <li>Metadata capture</li> <li>DDL generation</li> <li>Business rule documentation and validation</li> </ul>
Parsing and standardization	<ul style="list-style-type: none"> <li>Flexible definition of patterns and rules for parsing</li> <li>Flexible definition of rules for transformation</li> <li>Knowledge base of known patterns</li> <li>Ability to support multiple data concepts (individual, business, product, etc.)</li> <li>Manageable transformation actions</li> </ul>
Identity resolution	<ul style="list-style-type: none"> <li>Entity identification</li> <li>Record matching</li> <li>Record linkage</li> <li>Record merging and consolidation</li> <li>Flexible definition of business rules</li> <li>Knowledge base of rules and patterns</li> <li>Integration with parsing and standardization tools</li> <li>Advanced algorithms for deterministic or probabilistic matching</li> </ul>
Cleansing and enhancement	<ul style="list-style-type: none"> <li>Flexible definition of cleansing rules</li> <li>Knowledge base of common patterns (for cleansing)</li> <li>Knowledge base of enhancements (e.g., address cleansing, geocoding)</li> </ul>
Auditing and monitoring	<ul style="list-style-type: none"> <li>Data validation</li> <li>Data controls</li> <li>Services-oriented</li> <li>Rule management</li> <li>Rule-based monitoring</li> <li>Reporting</li> </ul>

- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## CONCLUSION

Even though many of the more established data quality tool vendors have been acquired by even bigger fish, there are still companies emerging with better approaches to fill the void. Whether better algorithms packaged in a different way, improvements in performance, better suitability to SOA, or even an open source offering, there is a wide range of vendors, products and tools to fit almost any organization's needs. Even armed with the knowledge of what you should look for in data quality tools, there is one last caveat: If your organization has an opportunity for data quality improvement, make sure that you have done your homework in business needs assessment and development of a reasonable RFP before evaluating and purchasing tools.

---

### ABOUT THE AUTHOR



**David Loshin**, president of [Knowledge Integrity, Inc.](#) is a recognized thought leader and expert consultant in the areas of data governance, data quality methods, tools, and techniques, master data management, and business intelligence. David is a prolific author regarding BI best practices, either via his B-Eye Network expert channel and numerous books on BI and data quality. His book, *Master Data Management*, has been endorsed by data management industry leaders, and his valuable MDM insights can be reviewed at [www.mdmbook.com](http://www.mdmbook.com). David can be reached at [loshin@knowledge-integrity.com](mailto:loshin@knowledge-integrity.com), or 301-754-6350.



[INTRO](#)

[DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)

[INDEX](#)

[BUSINESS OBJECTS \(NOW SAP\)](#)

[DATACTICS](#)

[DATAFLUX](#)

[DATALEVER](#)

[DATA MENTORS INC.](#)

[DATANOMIC](#)

[EPRENTISE](#)

[FUZZY! INFORMATIK \(NOW SAP\)](#)

[HARTE-HANKS TRILLIUM SOFTWARE](#)

[HUMAN INFERENCE](#)

[IBM](#)

[INFORMATICA](#)

[INNOVATIVE SYSTEMS INC.](#)

[NETRICS](#)

[ORACLE CORP](#)

[PERVASIVE SOFTWARE](#)

[PITNEY BOWES GROUP 1](#)

[STALWORTH INC.](#)

[TALEND](#)

[UNISERV](#)

[ZOOMIX \(NOW MICROSOFT\)](#)

[ABOUT SEARCHDATA-MANAGEMENT.COM](#)

[METHODOLOGY](#)

# Index at a Glance

Click on the product name at left to jump to a longer description.

PAGE	VENDOR	PRODUCTS	SAAS OR SERVICES	ON PREMISE SW
12	Business Objects (now SAP)	BusinessObjects Data Quality, Universal Data Cleanse, BusinessObjects Data Quality for SAP, BusinessObjects Data Quality for Oracle's Siebel CRM, BusinessObjects Data Insight		◆
12	Datactics	DataTrawler		◆
13	DataFlux	DataFlux Data Quality Integration Platform		◆
13	DataLever	DataLever Enterprise Suite		◆
14	Data Mentors Inc.	DataFuse	●	◆
14	Datanomic	dn:Director		◆
15	eprentise	eprentise Data Quality		◆
15	Fuzzy! Informatik (now SAP)	Fuzzy! Dime, Fuzzy! Analyzer, Fuzzy! Double, Fuzzy! Move, Fuzzy! Bank, Fuzzy! Umzug		◆
16	Harte-Hanks Trillium Software	Trillium Software System: TS Insight, TS Discovery, TS Quality, TS Enrichment	●	◆
16	Human Inference	Hlquality Suite	●	◆
17	IBM	IBM Information Server WebSphere QualityStage		◆
17	Informatica	Informatica Data Quality		◆
18	Innovative Systems Inc.	i/Lytics Enterprise Data Quality Suite: i/Lytics Data Profiler, i/Lytics Data Quality, i/Lytics GLOBAL	●	◆
18	Netrics	Netrics Matching Platform		◆
19	Oracle Corp	Data Quality for Oracle Data Integrator		◆
19	Pervasive Software	Pervasive Data Profiler		◆
20	Pitney Bowes Group 1	Customer Data Quality Platform, CDQ On Demand, Global Sentry, CDQ Platform for Microsoft Dynamics CRM, CDQ Platform for Salesforce.com, CDQ Platform for SAP and CDQ Platform for Siebel	●	◆
20	Stalworth Inc.	DQ*Plus, DQ*Plus Batch and Persistent, DQ*Plus Interactive	●	◆
21	Talend	Talend Open Profiler		◆
21	Uniserv	Data Quality Batch Suite	●	◆
22	Zoomix (now Microsoft)	Zoomix Accelerator		◆

**Vendor:** Vendor/developer of product at directory press time; ● **SaaS or services** indicates technology available as SaaS, hosted, on-demand, ASP and web services; ◆ **On-premise SW** indicates software or systems on premise; Ⓞ Description was written by the SearchDataManagement.com editorial team based on information gathered from vendor websites.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## BUSINESS OBJECTS, AN SAP CO.

- [BusinessObjects Data Quality](#)
- [Universal Data Cleanse](#)
- [BusinessObjects Data Quality for SAP](#)
- [BusinessObjects Data Quality for Oracle's Siebel CRM](#)
- [BusinessObjects Data Insight](#)

*Business Objects has multiple products for data quality monitoring, analyzing, standardizing and reporting. ♦*

**COMPANY WEBSITE:** [www.businessobjects.com](http://www.businessobjects.com)

**FOUNDED:** 1990

**SUMMARY:** BusinessObjects Data Quality standardizes, corrects, enhances and unifies data from various sources. Universal Data Cleanse parses, standardizes and identifies non-customer and region-specific data. BusinessObjects Data Quality for SAP is embedded within SAP for global address correction and verification. BusinessObjects Data Quality for Oracle's Siebel CRM is similarly embedded within the Siebel CRM. BusinessObjects Data Insight supports monitoring, analyzing and reporting data quality in a relational database or a flat file in an open system or mainframe environment. ◉

**PRICING:** Declined to provide pricing.

## DATACTICS

### DataTrawler

*Dataactics DataTrawler is data quality software with grid technology for large data volumes. ♦*

**COMPANY WEBSITE:** [www.dataactics.com](http://www.dataactics.com)

**FOUNDED:** 1999

**SUMMARY:** DataTrawler combines data processing techniques with the latest generation of data management technologies. It is based on "fuzzy" technology, namely the tolerance of errors per length of any given string of data, and provides users with dials to adjust the level of errors that they're prepared to tolerate. DataTrawler uses next-generation grid technology to call on available computing power and allow processes to be shared across many computers. Dataactics also has a "data management methodology" with three phases: analyze, re-engineer and match—and three management elements: report, integrate and manage.

**PRICING:** Declined to provide pricing.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## DATAFLUX CORP.

### DataFlux Data Quality Integration Platform

*DataFlux has multiple products supporting different data quality processes, all based on a single platform. ♦*

**COMPANY WEBSITE:** [www.dataflux.com](http://www.dataflux.com)

**FOUNDED:** 1997

**SUMMARY:** DataFlux data quality integration products have functions for organizations to: plan and prioritize data-correction programs; parse data into components to help identify and resolve problematic data; standardize, correct and normalize data to create a unified view of corporate information; verify and validate data accuracy to improve the overall accuracy of customer records, product data and other information; and apply business rules across the enterprise to ensure that all corporate data reflects business needs. All DataFlux products are built from the same core technology and code base. ◉

**PRICING:**

DataFlux dfPower Studio is available on a per user basis, while the DataFlux Integration Server utilizes a per server model. The pricing of the DataFlux Data Quality Integration Platform starts at \$50,000 and varies according to the number of users and the processing speed of the servers involved.

## DATALEVER

### DataLever Enterprise Suite

*DataLever Enterprise Suite is a visual process designer for creating and running high-performance data-transformation and data quality processes. ♦*

**COMPANY WEBSITE:** [www.datalever.com](http://www.datalever.com)

**FOUNDED:** 1998

**SUMMARY:** DataLever Enterprise Suite is a visual process designer allowing organizations to create and run high-performance data-transformation processes. Using a modular approach, components are selected from a palette. Components perform specialized operations. By connecting the components, custom data processing engines, which resemble flow charts, can be created. The results of the job can be stored in a database, viewed in a spreadsheet or written to a report. The DataLever framework encompasses a broad range of functionality, including comprehensive data re-engineering and data profiling. ◉

**PRICING:** Depending on database size, pricing begins at \$8,000 per month for ASP-delivered functions.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## DATAMENTORS, INC.

### DataFuse

*DataFuse is a modular database preparation and integration system that cleanses, organizes, standardizes, matches and prepares data.* ● ◆

**COMPANY WEBSITE:** [www.datamentors.com](http://www.datamentors.com)

**FOUNDED:** 1998

**SUMMARY:** DataMentors provides database preparation and maintenance marketing systems featuring DataFuse (which cleanses, organizes, standardizes, matches and prepares data), ValiData and PinPoint. These may be used individually or as an integrated end-to-end turnkey database system for a clean, granular and complete view of the customer through data validation, transformation, standardization and householding. Offered as either a customer-premise installation or ASP-delivered system, DataMentors supports proprietary data discovery, reporting and analysis, campaign management, data mining, and modeling practices.

**PRICING:** The pricing structure is based on several factors including, but not limited to: number of licenses, database size, length of term, specific client product customization and client support criteria. Information as of Q1 2009. Declined to provide additional information.

## DATANOMIC

### dn:Director

*dn:Director is a multi-function data quality platform.* ◆

**COMPANY WEBSITE:** [www.datanomic.com](http://www.datanomic.com)

**FOUNDED:** 2001

**SUMMARY:** Datanomic's dn:Director is a multi-function data quality platform, providing a range of data quality processors for profiling, checking, transforming, parsing, matching, enhancing, merging and reporting on data from disparate systems, languages and standards, accessible from a single interface. dn:Director supports a data-led discovery methodology where data sources are examined for their content and inherent business rules. Deployments might include data migrations, linking records across disparate sources, single view or master data creation and compliance solutions. dn:Director is an all-Java application.

**PRICING:** Datanomic's dn:Director offers a flexible pricing model for different deployments, with list prices starting at \$20,000. Modules of the product (e.g., profiling, audit, transformation, text analysis, matching) may be purchased separately. Licenses may also be priced by data volume, hardware, users or software purpose.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIK \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## EPRENTISE

### eprentise Data Quality

*eprentise Data Quality is software that standardizes, locates and resolves duplicate data. ♦*

**COMPANY WEBSITE:** [www.eprentise.com](http://www.eprentise.com)

**FOUNDED:** 2006

**SUMMARY:** eprentise Data Quality software applies standard abbreviations and naming conventions to enterprise data with find-and-replace rules - which can be used to change punctuation and abbreviations or make global text replacements. User-defined criteria and complex Boolean rule structures identify candidate duplicates and operate across tables. The software compares records and identifies candidate duplicates based on matches to these criteria. Matches are assembled into duplicate sets for the user to verify and resolve. Software is currently available on Oracle e-Business Suites.

**PRICING:** eprentise Data Quality software starts at \$150,000.

## FUZZY! INFORMATIK, AN SAP CO.

### Fuzzy! Dime, Fuzzy! Analyzer, Fuzzy! Double, Fuzzy! Move, Fuzzy! Bank, Fuzzy! Umzug

*Fuzzy! Informatik, an SAP Co., has data quality products primarily for international business partner data. ♦*

**COMPANY WEBSITE:** [www.fazi.de/index.php?en](http://www.fazi.de/index.php?en)

**FOUNDED:** 1994

**SUMMARY:** Fuzzy! products, available worldwide, are designed for international business partner data management. Data quality products include: Fuzzy! Dime (monitoring and measuring data quality); Fuzzy! Analyzer (structuring and analyzing information); Fuzzy! Double (preventing duplicates and fault-tolerant searches); Fuzzy! Post (qualifying addresses); Fuzzy! Move (searching for customers who have moved to unknown destinations); Fuzzy! Bank (checking and correcting bank details); Fuzzy! Tel (finding current telephone numbers); and Fuzzy! Umzug (updating customer databases with addresses of customers who have moved). ◉

**PRICING:** Declined to provide pricing.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## HARTE-HANKS TRILLIUM SOFTWARE

### **Trillium Software System: TS Insight, TS Discovery, TS Quality, TS Enrichment**

*Trillium Software System is an integrated global data quality management software suite. ● ◆*

**COMPANY WEBSITE:**

[www.trilliumsoftware.com](http://www.trilliumsoftware.com)

**FOUNDED:** 1993

**SUMMARY:** Trillium Software System: Integrated suite of four data quality products, providing data quality detection, construction, repair and maintenance. TS Insight is a browser-based window into data and a data quality dashboard for business users/IT professionals to manage data. TS Discovery is an automated data profiling tool that analyzes data to reveal weaknesses and gaps. TS Quality cleanses data from different sources. TS Enrichment supplements business/corporate data with demographic and geographic information to improve postal performance and more. Some products available as on-premise software or on-demand.

**PRICING:** Declined to provide pricing.

## HUMAN INFERENCE

### **Hlquality Suite**

*Hlquality Suite is a data quality management platform based on natural language processing. ● ◆*

**COMPANY WEBSITE:**

[www.humaninference.com](http://www.humaninference.com)

**FOUNDED:** 1986

**SUMMARY:** Human Inference's open data quality software aims to help organizations optimize customer databases for increased revenue, reduced cost, improved decision-making and regulatory compliance. The Hlquality Product Suite is applied in call center solutions, data warehouses, customer relationship management (CRM), and sales and marketing systems. It is used in de-duplication and matching processes, and for fraud detection and centralizing relationship databases. Human Inference has data quality products available as on-premise software or, via partnerships, as SaaS (may have limited availability). ○

**PRICING:** License fees are subject to number of records in customer database(s). The average starting price of the products is approximately €30,000 (approximately \$38,361 USD at press time).





- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATAEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## IBM

### **IBM Information Server WebSphere QualityStage**

*IBM Information Server WebSphere QualityStage provides data quality management for customer and business partner address information. ♦*

**COMPANY WEBSITE:** [www.ibm.com/us](http://www.ibm.com/us)

**FOUNDED:** 1924

**SUMMARY:** IBM Information Server WebSphere QualityStage modules are designed to support data quality management for customer and business partner address information, with: address verification and enrichment modules; address certification, including the US Postal Service, Coding Accuracy Support Systems (CASS) and more, which supports verification and enrichment of addresses; Geospatial information to pinpoint locations and establish spatial; and certified data quality integration for the SAP Business Address Services DES, which provides extensible search and match capabilities to SAP customers. ◉

**PRICING:** IBM WebSphere QualityStage for EIM 50 Concurrent Users License, plus software subscription and support for 12 months (D59B1LL), is \$25,000 (According to IBM's website as of Q2 2008).

## INFORMATICA

### **Informatica Data Quality**

*Informatica Data Quality has data analysis, cleansing, matching, reporting and monitoring capabilities to implement and manage enterprise-wide data quality initiatives. ♦*

**COMPANY WEBSITE:** [www.informatica.com/Pages/index.aspx](http://www.informatica.com/Pages/index.aspx)

**FOUNDED:** 1993

**SUMMARY:** Informatica Data Quality has data analysis, cleansing, matching, reporting and monitoring capabilities. It addresses master data types, including customer, product, financial, materials, pricing, order and asset. Features include a data quality workbench (design, build and manage data quality programs, create reports and dashboards, deploy data quality rules in real time), data quality assistant (edit, review low-quality records, track changes for auditing), data quality profiling, data cleansing and parsing, data matching, scalable deployment and a global component software development toolkit and more. ◉

**PRICING:** Declined to provide pricing.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## INNOVATIVE SYSTEMS INC.

### **i/Lytics Enterprise Data Quality Suite: i/Lytics Data Profiler, i/Lytics Data Quality, i/Lytics GLOBAL**

*i/Lytics Enterprise Data Quality Suite is a data quality platform with functions for data profiling and analysis, data quality and address validation. ●◆*

**COMPANY WEBSITE:** [www.innovativesystems.com](http://www.innovativesystems.com)

**FOUNDED:** 1968  
**SUMMARY:** Innovative Systems' i/Lytics Enterprise Data Quality Suite includes: i/Lytics Data Profiler for data profiling and analysis; i/Lytics Data Quality for data cleansing, data linking and change management; and i/Lytics GLOBAL, a CASS-certified address validation and Geocoding product. The i/Lytics Enterprise Data Quality Suite offers multiple deployment styles, available as licensed software in both real-time and batch mode, as a hosted ASP solution or on a service bureau basis. Innovative Systems has data quality products and services available as on-premise software or on-demand.

**PRICING:** Declined to provide pricing.

## NETRICS

### **Netrics Matching Platform, Netrics Matching Engine, Netrics Decision Engine, Netrics Reporting Engine, Netrics N-Mend Data Reconciliation Tool**

*Netrics Matching Platform is designed for finding and linking data. ◆*

**COMPANY WEBSITE:** [www.netrics.com](http://www.netrics.com)  
**FOUNDED:** 2000

**SUMMARY:** Netrics Matching Platform includes: Netrics Matching Engine, which matches data with error tolerance that approximates human perception; and Netrics Decision Engine, which creates automated decision-models for detecting duplicates, linking records, resolving entities and more. Netrics creates custom models for the specific datasets, market conditions and business requirements of an application. Netrics Reporting Engine gives details on the state of data, in a data-quality reporting tool. Netrics N-Mend Data Reconciliation Tool is a Web-based tool that locates/eliminates duplicates from databases and more.

**PRICING:** The Netrics Matching Platform and the Netrics Decision Engine each cost approximately \$25,000 per CPU-core.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## ORACLE CORP.

### Data Quality for Oracle Data Integrator

Oracle Data Quality for Oracle Data Integrator includes multiple tools for data quality and data governance. ♦

**COMPANY WEBSITE:** [www.oracle.com](http://www.oracle.com)

**FOUNDED:** 1997

**SUMMARY:** Oracle Data Quality for Oracle Data Integrator is a data quality platform that covers various data quality needs. Its rule-based engine and scalable architecture support data quality and name and address cleansing. Oracle Data Quality enables global data quality support, name and address cleansing, parsing and standardization, customer data validation with postal directory or third-party information, automatic process duplication identity, user-customizable rules, customizable data quality process and rules, built-in name and address standardization validation, and enrichment.

**PRICING:** Data Quality for Data Integrator (up to a maximum of 100 million records), Processor License: \$60,000; Software Update License and Support: \$13,200. Data Quality Rules for Data Integrator (per rule set): License Price: \$20,000; Software Update License and Support: \$4,400 (According to Oracle's website as of Q2 2008).

## PERVASIVE SOFTWARE

### Pervasive Data Profiler

Pervasive Data Profiler automates the testing, quality control and regulatory compliance of critical data. ♦

**COMPANY WEBSITE:** [www.pervasive.com](http://www.pervasive.com)

**FOUNDED:** 1994

**SUMMARY:** Pervasive Data Profiler helps ensure data quality by enabling users to audit various types of data and automate testing against changing business needs and compliance regulations. Users can assess data across multiple data platforms and quarantine questionable data until it can be cleansed. Pervasive Data Profiler also supports the testing of large amounts of transactional data.

**PRICING:** Pricing starts at \$5,000.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATAEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## PITNEY BOWES GROUP 1 SOFTWARE

- [\*\*Customer Data Quality Platform\*\*](#)
- [\*\*CDQ On Demand\*\*](#)
- [\*\*CDQ Global Sentry\*\*](#)
- [\*\*CDQ Platform for Microsoft Dynamics CRM\*\*](#)
- [\*\*CDQ Platform for Salesforce.com\*\*](#)
- [\*\*CDQ Platform for SAP\*\*](#)
- [\*\*CDQ Platform for Siebel\*\*](#)

*Pitney Bowes Group 1 is a provider of customer data quality and data profiling software. ● ◆*

**COMPANY WEBSITE:** [www.g1.com](http://www.g1.com)

**FOUNDED:** 1982

**SUMMARY:** Pitney Bowes Group 1 data quality software creates a view of customer relationships to support targeting, generate communications and more. The vendor features a customer data quality suite -- Customer Data Quality Platform, CDQ On Demand, Global Sentry, CDQ Platform for Microsoft Dynamics CRM, CDQ Platform for Salesforce.com, CDQ Platform for SAP and CDQ Platform for Siebel -- as well as various standalone applications (Merge/Purge Plus, Business Merge/Purge Plus, Dispatcher 4, Generalized Selection Plus, List Conversion Plus and EZ-Case Plus). The vendor has data quality products and services available as on-premise software or services. ○

**PRICING:** Declined to provide pricing.

## STALWORTH INC.

### **DQ\*Plus**

*DQ\*Plus is a data quality platform with functions for data cleansing and consolidation for databases and enterprise applications. ● ◆*

**COMPANY WEBSITE:** [www.stalworth.com](http://www.stalworth.com)

**FOUNDED:** 1998

**SUMMARY:** Stalworth Inc. and Melissa Data formed a strategic partnership to deliver DQ\*Plus.. DQ\*Plus Batch and Persistent cleans data automatically, while DQ\*Plus Interactive works with users inside their applications. DQ\*Plus has a service-oriented architecture (SOA) and a web services API that supports interoperability with legacy and third-party applications. Application-aware connectors, packaged rules, a point-and-click interface and DQ\*Plus SIMULATOR enable businesses to deploy DQ\*Plus and enable business users to configure data quality rules. Stalworth has data quality products and services available as on-premise software or as web services. ○

**PRICING:** Pricing for the DQ\*Plus Enterprise Suite begins at \$50,000.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATIC \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## TALEND

### Talend Open Profiler

*Talend Open Profiler is an open source data profiler. ♦*

**COMPANY WEBSITE:** [www.talend.com](http://www.talend.com)

**FOUNDED:** 2006

**SUMMARY:** Talend Open Profiler, an open source data profiler, enables companies to assess the quality of data and decide which actions must be taken to correct "dirty data" can negatively affect an organization. Talend Open Profiler's features include: metadata discovery, which identifies the structure of the databases that need to be analyzed; statistics definition, which defines the statistics and metrics that need to be measured on each data item; and results and graphs, which make it easy to view the results and assess the level of quality of the data.

**PRICING:** Talend Open Profiler is open source, available at no cost under a GPL license at [www.talend.com](http://www.talend.com).

## UNISERV, GMBH

### Data Quality Batch Suite

*Uniserv provides data quality software and services for customer data, particularly for business intelligence and customer relationship management applications. ● ♦*

**COMPANY WEBSITE:** [www.uniserv.com](http://www.uniserv.com)

**FOUNDED:** 1969

**SUMMARY:** Uniserv, GmbH is a German supplier of data quality software and services for the quality assurance of customer data in the areas of business intelligence (BI), customer relationship management (CRM) applications, data warehousing, eBusiness and direct and database marketing. Uniserv's data quality products feature: address analysis and formatting; merge, purge and duplicate check; intelligent search and duplicate check; postal address check; geo- and micromarketing; postage and mailing optimization; relocation addresses; direct marketing suite; banking data finder; and telephone number finder. ◎

**PRICING:** Declined to provide pricing.



- [INTRO](#)
- [DATA QUALITY MANAGEMENT AND CHOOSING DATA QUALITY TOOLS](#)
- [INDEX](#)
- [BUSINESS OBJECTS \(NOW SAP\)](#)
- [DATACTICS](#)
- [DATAFLUX](#)
- [DATALEVER](#)
- [DATA MENTORS INC.](#)
- [DATANOMIC](#)
- [EPRENTISE](#)
- [FUZZY! INFORMATICA \(NOW SAP\)](#)
- [HARTE-HANKS TRILLIUM SOFTWARE](#)
- [HUMAN INFERENCE](#)
- [IBM](#)
- [INFORMATICA](#)
- [INNOVATIVE SYSTEMS INC.](#)
- [NETRICS](#)
- [ORACLE CORP](#)
- [PERVASIVE SOFTWARE](#)
- [PITNEY BOWES GROUP 1](#)
- [STALWORTH INC.](#)
- [TALEND](#)
- [UNISERV](#)
- [ZOOMIX \(NOW MICROSOFT\)](#)
- [ABOUT SEARCHDATA-MANAGEMENT.COM](#)
- [METHODOLOGY](#)

## ZOOMIX (MICROSOFT SUBSIDIARY)

### Zoomix Accelerator

*Zoomix Accelerator is a data processing server designed to solve data inconsistencies in-line with business processes. ♦*

**COMPANY WEBSITE:** [www.zoomix.com](http://www.zoomix.com)

**FOUNDED:** 1999

**SUMMARY:** Zoomix, a Microsoft Subsidiary, develops and markets software to support the implementation of master data management (MDM) systems. The company's technology combines semantic analysis with machine learning to classify, match and standardize corporate master data (including semi-structured, highly variable product data and financial data). Zoomix Accelerator aims to help organizations resolve inconsistencies in data at the source and in-line with routine business workflows. ◉

**PRICING:** Declined to provide pricing.



**SearchDataManagement.com** is a guide for data management professionals and business leaders. With its combination of news, learning guides, expert advice, white papers, webcasts and customized research, SearchDataManagement.com offers a rich collection of insight on how to efficiently manage the data supply chain. The site also offers tips on vendors and product selection, as well as expert advice.

Visit SearchDataManagement.com for:

- ▶ **Independent content:** Award-winning, vendor-independent news and analysis.
- ▶ **Expert advice:** Our Ask the Experts section features advice from some of the leading authorities in the data management domain.
- ▶ **Learning guides:** Search our comprehensive library for useful how-to guides on any data management topic.
- ▶ **Vendor-produced content:** We hand-select white papers and webcasts to address the most relevant data management trends, issues and solutions.

---

**Guide methodology:** TechTarget has not evaluated the products listed &/or described in this Directory and does not assume any liability arising out of the purchase or use of any product described herein, neither does it convey any license or rights in or to any of the evaluated or listed products. TechTarget has prepared this Directory from sources deemed reliable (including vendors, research reports and certain publicly available information). TechTarget has used good faith efforts to indicate when content has been provided by a vendor and, in some cases, has removed what it has deemed to be overt marketing language.

TechTarget is not responsible for any errors or omissions contained in this Directory or for interpretations thereof, and expressly disclaims all warranties as to the accuracy, completeness or adequacy of all content contained herein. This disclaimer of warranty is in lieu of all warranties whether expressed, implied or statutory, including implied warranties of merchantability or fitness for a particular purpose. Information in this Directory is current as of June 30, 2008. SearchDataManagement.com editors contacted all vendors for updates on pricing and product information in Q1 2009. Any updates vendors supplied were incorporated into the 2009 edition of the directory. For more recent information, please check the vendor's websites. The opinions expressed herein are subject to change without notice.

©2009 TechTarget, Inc. All Rights Reserved. Reproduction and distribution of this publication in any form without prior written permission is forbidden. TechTarget and the TechTarget logo are registered trademarks of TechTarget, Inc.; all other trademarks are the property of their respective companies.

To compile this guide, our editorial team initially consulted research reports by major analyst firms covering the data quality market and contacted vendors about products reviewed by those firms. Editors also conducted additional Internet research and solicited feedback from our expert contacts. A notice about the project was posted on SearchDataManagement.com and listed regularly in our email newsletters.

Vendors were invited to submit listings via a form on the website. For vendors that did not submit listings, our editorial team compiled listings by excerpting information from the vendor's website. All entries, whether they were vendor-submitted or compiled by our team, were edited for length and clarity and to remove overt marketing language. In order to best assist our readers in assessing products, our editorial team attempted to obtain basic pricing information for all products in this directory—requesting information from vendors multiple times via email. Vendors that did not respond, or refused to provide any pricing information, have this statement on their listings: "Declined to provide pricing."

Collection of data for this directory took place during the second calendar quarter of 2008. As with any directory of this kind, products and vendors may change substantially at any time. Though every effort was made to make this directory as complete and accurate as possible, there may be changes, errors, omissions or vendors in this market not included in this guide. Nothing in this guide should be construed as endorsements, professional suggestions or advice. This directory should be used simply as a resource. We strongly urge you to supplement this with your own research and to contact vendors for the most up to date information about their companies or products. It is our intent to update this directory annually, but that is subject to change.