

Taming Data Chaos: Simplify Data Integration with DataDirect Data Integration Suite

An ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) White Paper
Prepared for DataDirect

March 2009



IT MANAGEMENT RESEARCH,
INDUSTRY ANALYSIS AND CONSULTING

Table of Contents

Introduction	1
Business Applications and the Challenge of Heterogeneous Data.....	1
DataDirect Data Integration Suite and Where it Fits	3
Customer Anecdotes and Use Cases	5
ECN Group	5
DataFlux	6
Core Computer Group	6
DataDirect Differentiators, Benefits, and Value Proposition	7
EMA Perspective	9
About DataDirect Technologies	9

Introduction

Each year, the role of Information Technology becomes more complex. Virtually every IT organization of any size follows the latest trends, introducing new platforms, applications and infrastructure on an ongoing basis. At the same time, most continue to support applications and platforms that may be decades old. Technologies across this wide spectrum – from Rich Internet Applications (RIAs) to mainframes, from .NET to COBOL, and from Web Services deployments to legacy Electronic Data Interchange (EDI) – coexist within today’s data centers, each one adding unique challenges to the overall task of IT support.

Valuable business information is locked within these diverse systems, and IT organizations are confronting the twin problems of accessing and integrating data. Extensible Markup Language (XML) has come into its own as the “lingua franca” – the de facto language – of data exchange. Companies that have leveraged EDI for decades are suddenly facing the need to integrate with partners and suppliers using XML-based technologies. This represents a paradigm shift for businesses and development teams alike. Extracting and updating business-relevant data contained within such documents presents challenges that include training, performance, and data integrity considerations.

At the same time, the vast majority of IT organizations are struggling with funding and facing the need to balance dwindling budgets against rising costs. Leading companies are exploring a variety of ways to deal with these challenges and to do so cost-effectively. Taking a cue from best practice disciplines such as the IT Infrastructure Library (ITIL) and Control Objectives for IT (COBIT), many are turning to automation to work “smarter.” Data transformation and integration are key areas where specialized products can yield impressive time and cost efficiencies.

The DataDirect Data Integration Suite is one of the leaders in this space. By easing the pain of integrating diverse applications, the Suite delivers time, cost, and performance efficiencies. A lightweight, cost-effective solution, it provides easy access to data stored in virtually any format – including Relational Database (RDBMS) files, XML, Web Services, Microsoft Excel/Office, flat files, and EDI documents. In doing so, it leverages standards-based XML and XQuery technologies, transforming the data contained within these diverse data stores into XML-compatible format.

Alone or in conjunction with other DataDirect solutions, such as Shadow (for mainframe access), the Data Integration Suite can help reverse rising costs via savings in multiple areas. This paper describes the Suite and its capabilities, and illustrates its value proposition with a variety of customer testimonials and use cases across diverse industries.

Business Applications and the Challenge of Heterogeneous Data

Integration – with partners, customers, and suppliers – is the lifeblood of many of today’s businesses. In recent ENTERPRISE MANAGEMENT ASSOCIATES® (EMA™) research, 55% of IT organizations surveyed reported business to business (B2B) integrations, 52% reported Service Oriented Architecture (SOA) and/or Web Services deployments, and 49% reported Enterprise Application Integration (EAI) applications. Development teams are hard at work on such projects, with 78% of companies reporting support for custom applications within the enterprise – the highest of any application type surveyed.

Under the covers, virtually every integration project is really about data – sending, receiving, and transforming data across applications, companies, and customers. Supply chain applications exchange order information with suppliers and billing information with customers. The complexity surrounding such seemingly simple activities can be overwhelming, and a typical EDI integration, for example, takes months to complete. Every company’s data is different, and, because of ongoing mergers and acquisitions, many companies are actually managing and integrating multiple internal data sets.

In today’s companies, data can exist on disparate platforms (UNIX, Windows, mainframe), in dissimilar Relational Databases (SQL Server, Oracle, MYSQL), and in multiple formats (Microsoft Excel spreadsheets/Office documents, proprietary vendor files). With little or no consistency between businesses, or even between departmental data within a single company, virtually every integration is unique and requires custom programming (See ECN Group Case Study below). Without specialized products, combining data stored in multiple formats – for example, mainframe data with data stored in an Excel spreadsheet – requires specialized skills and custom code. Integrating unlike platforms, data stores, and data formats is a complex business, and this complexity is one factor driving today’s high development and support costs.

For many medium-sized to enterprise-sized businesses, application heterogeneity is also a factor contributing to “data chaos” (See Figure 1). Diverse data, stored in multiple locations and on multiple platforms, can be accessed and consumed by Java and .NET applications, SOA services, REST and Scripting applications, and legacy software. This diversity takes its toll. For developers and architects, it requires ongoing training as well as substantial manual coding. It wreaks havoc with governance and data compliance (GRC) initiatives, since data is accessed by multiple technologies using multiple methods. It also has performance implications, since it is impossible to predict which applications are requesting the same data concurrently. Finally, it drives up development costs. Data schemas constantly change to meet the changing needs of the business, and multiple programs must be modified with each change to a data source.

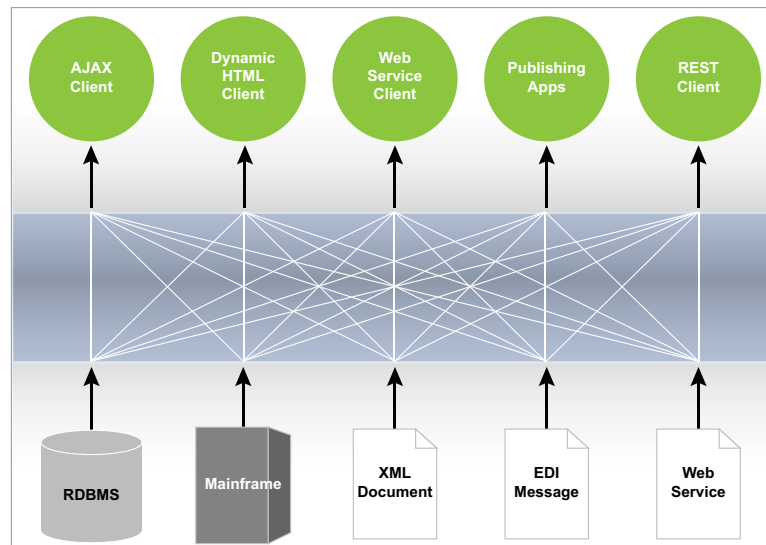


Figure 1: Data Chaos
Multiple technologies accessing diverse data sources with diverse protocols

In short, while applications built on these technologies promise high value and competitive differentiation, they do not come cheaply. Transforming data into business-focused information – messaging it into a form that is useful to the business – is an expensive proposition.

Customer Quote: “An implementation that would have taken us 24 hours the old way should now take us about 12. So we’re talking about at least a 50 percent reduction in implementation cost.”

DataDirect Data Integration Suite and Where it Fits

The DataDirect Data Integration Suite is specifically designed to reduce the impact of data and platform heterogeneity. Instead of learning a different Application Program Interface (API) or protocol for every platform, developers rely on XQuery statements “talking” to the Data Integration Suite, which communicates directly with the data source. XQuery is the industry standard language for accessing data stored within XML documents, in the same way that Structured Query Language (SQL) is utilized to query data stored in relational systems. With DataDirect in place, developers use a standard set of commands to access data, regardless of the platform on which it is stored.

In effect, the Data Integration Suite “translates” data from different platforms and formats it into standardized Extended Markup Language, or XML (See Figure 2). Broad technology coverage (See Table 1) enables a significant portion of the work of data transformation and integration typically performed by developers and architects to be done via intelligent automation.

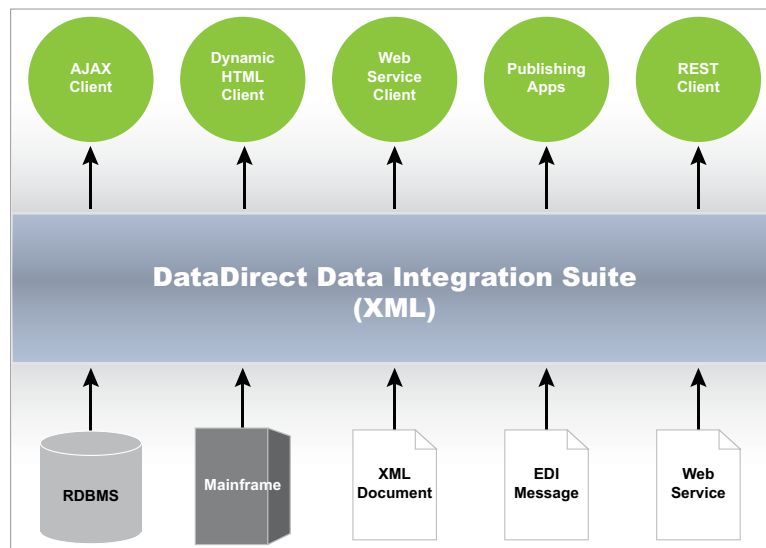


Figure 2: Unraveling Data Chaos
Multiple technologies “speaking a common language”

These capabilities can deliver impressive time and cost efficiencies. With pre-defined converters, developers can generate schemas in minutes using a step-by-step wizard. This approach replaces custom coding, which can take several days or longer for a single application. Developers and architects can focus on solving problems instead of on writing repetitive, low-level data access routines.

The suite consists of three bundled, integrated solutions:

- **Stylus Studio®** is the graphical portal that developers use to do their work. It is an Integrated Development Environment (IDE) that simplifies data integration with drag-and drop mapping, code generation, and debugging for Java, C#, and .NET. It supports multiple XML technologies including XQuery, XSL Transformations (XSLT), Web Services, and XML Pipelines. Recent enhancements support conversion of EDI to XML, XQuery Document, and XML Schema Definition (XSD) Documents.

- For a complete description visit: http://www.stylusstudio.com/xml_brochure.html

- **DataDirect XQuery®** is an extension of the World Wide Web Consortium (W3C) standard XQuery language. Since XML is “shaped” like an outline instead of in tabular rows and columns like relational data, standard Structured Query Language statements cannot efficiently extract business-relevant data embedded within XML documents. A different access methodology is necessary, and XQuery fills this gap.

In conjunction with DataDirect XML Converters, DataDirect XQuery simplifies data access and exchange for virtually any data source. It can also integrate data from multiple diverse sources into a single result for XML-based data interchange. DataDirect XQuery supports the XQuery update standard for XML (XUF), Web Services, streaming, and document projection.

- Additional information is available at: <http://www.xquery.com/products/factsheet.html>

- **DataDirect XML Converters®** are high performance components providing conversion and bidirectional data exchange between XML and virtually any data file format including EDI, flat files, and custom formats. They are the “core” of the DataDirect Data Integration Suite, and include both Java and .NET converters. Recent enhancements support Health Insurance Portability and Accountability Act (HIPAA) version 5010 and full integration with Microsoft BizTalk.

- For more information, visit: <http://www.xmlconverters.com/products/factsheet.html>

Broad Technology Coverage
Platforms and Technologies: Oracle, Microsoft SQL Server, DB2, Sybase, MySQL, PostgreSQL, Informix, XML, Web Services (WSDL/SOAP, REST), OOXML & ODF documents
Bi-directional converters: EDIFACT, Health Level 7 (HL7), IATA, X12 (including HIPAA 5010, ICD-10), EANCOM, Edig@s, SEF (customizations for all EDI formats), Flat files (CSV, tab delimited, SYLK, dBase, JSON, and more)
Full list of versions at: xmlconverters.com/standards
Custom Conversions: Proprietary to XML (not bi-directional)

Table 1: DataDirect Data Integration Suite
Broad Technology Coverage

These solutions capitalize on DataDirect's years of experience developing solutions for RDBMS access, and build on industry standards. Optimally positioned for both integration projects and coding projects that rely on data from multiple sources, the Suite is designed for simple deployment and ease of use. It is well-positioned for solving project-level integration challenges with a low initial investment and requires virtually no consulting expertise to deploy.

The Suite also features other capabilities that streamline both coding and performance. For example, Business Data Models enable developers to group data requests from multiple heterogeneous sources into a single query. One call aggregates data into a single request, dramatically reducing both coding effort and execution overhead.

Customer Quote: "... We needed to integrate our product ordering Web Service with data from our Oracle and DB2 databases..."

Customer Anecdotes and Use Cases

ECN Group

Leading-edge IT organizations leverage automation to improve efficiency while reducing costs. The ECN group leveraged DataDirect Data Integration Suite to speed the onboarding process for new customers. Facing a business climate in which custom coding was cutting into profits, ECN reported a 50% reduction in implementation costs for each new customer.

Who: The ECN Group was initially established to provide a Managed Messaging Service for the New Zealand Customs Service. Over the past ten years, the company has expanded to offer a broad suite of integration and messaging solutions to a wide range of customers. Their primary role is as a messaging broker, a single point of control that centrally handles the complexity of document conversion across multiple companies within a customer's supply chain.

The problem: In a typical B2B transaction, a customer's trading partner sends ECN a file, such as an invoice, in its own internal format. ECN receives the file, converts it into the correct customer format, and then forwards it on to the customer for processing. This ensures that files can be processed successfully.

ECN's goal in turning to DataDirect was to increase profitability. Before onboarding a customer, ECN had to write code to convert from and to the correct file formats. Since every customer's systems were unique, this required significant custom coding for each new customer. This was not only expensive from the development perspective; it also delayed the process of converting a prospect into a paying customer—and collecting revenue. An ECN executive noted, "With our existing system we could not charge customers enough to cover our implementation costs, so we had to recoup it during production. We were making a lot of revenue but not as much profit as we'd like."

The solution: ECN evaluated multiple commercial B2B platforms; however, none totally satisfied their requirements. Shifting gears, they decided to leverage an XML tool set to create a solution in-house. They again evaluated multiple potential products, including DataDirect XML Converters and Stylus Studio. The deciding factor was service and support. "Service was the big differentiator. We liked that DataDirect would assign us one person entirely informed about our issues that we could get on the phone with. We were also very impressed with the information available on the support Web site."

The result: Although service was a decisive factor, the products themselves gave ECN the edge they needed. The ECN solution consists of a Java framework, with DataDirect XML Converters on the front end to convert the various incoming file formats to XML. ECN uses Stylus Studio for B2B translation. They estimate that, once all of the existing legacy implementations are converted over to the new platform, it will be handling approximately 25 million messages annually. "An implementation that would have taken us 24 hours the old way should now take us about 12. So we're talking about at least a 50 percent reduction in implementation cost."

ECN Group Case Study
Faster, cheaper customer implementations

DataFlux

DataFlux saw new market opportunities for integrating with emerging industry standards, such as Health Level 7 (HL7). It needed a tool that would help it take advantage of these opportunities, and turned to DataDirect. In an early implementation, it solved a scaling problem that had prevented a potential customer from loading a critical data file into a Data Warehouse application. The resulting solution, which enables the file to be processed in less than a minute, is now a high-demand item with other customers as well.

Who: DataFlux, a wholly owned subsidiary of SAS®, provides data quality and data integration technology. Its products enable companies to build a unified view of customer, product, and supplier data, or any other corporate data assets. DataFlux helps customers streamline data-driven initiatives via data integration, product information management, and compliance.

The problem: Multiple DataFlux customers rely on standards such as ACORD, HL7, and SWIFT. For DataFlux, these standards represent new market opportunities. However, the company lacked the technology to quickly implement and support these services.

The solution: A DataFlux customer wanted to use DataFlux's profiling and data monitoring technology to validate a data file before it was transferred from a legacy system to a data warehouse in XML format. Because the file was so large—600 MB—the customer's ODBC drivers were unable to process it. DataFlux turned to DataDirect for a solution.

The result: Using a standalone product that allowed DataFlux to transform the XML structure into multiple tables, and then feed them into the DataDirect XQuery engine, the file was processed in under a minute's time. Now, other DataFlux customers with similar problems plan to take advantage of these same capabilities.

DataFlux Case Study
Providing customers with flexible, efficient file processing

Core Computer Group

A third DataDirect customer, Core Computer Group, built a car crash information system over standards-based technology for the National Highway Traffic Safety Administration (NHTSA) and the Volpe National Transportation Systems Center (Volpe Center), both part of the U.S. Department of Transportation (DOT). It estimated that, using traditional Java coding and SQL data access, the project would require 50 Java classes and 150 SQL statements for each study type. Using DataDirect XQuery, it was able to store all data in XML format and create the same system capabilities using a single Java class and 7 XQueries.

Who: Core Computer Group, contractors to the National Highway Traffic Safety Administration (NHTSA).

The problem: The NHTSA required a system that delivered car crash information in an open format that followed U.S. standards and compliance rules. Previously, data distribution was done via flat files and a variety of proprietary formats. The engineering team investigated several options for the project, including the use of Java code and SQL. They estimated that converting data from a relational database into XML would require approximately 50 Java classes and 150 SQL statements for one study type.

The solution: After evaluating available options, the team decided that a better option would be to store data in an XML repository. By using DataDirect XQuery® to handle their XML and relational tasks, all of the sub-processing is performed on XML files. This approach required a single Java class and 7 XQueries.

The result: "Building the XML repository from the Oracle relational database and later executing the XQueries on the XML to traverse the data resulted in amazing performance," reports a Core Computer Group engineer. "A large extraction from 500MB of XML data, building datasets intended for SAS analysis, took significantly less time compared to using different methods."

Core Computer Group Case Study
Reduced 50 Java classes to 1, 150 SQL Statements to 7 XQueries

DataDirect is in the enviable position of having multiple customer testimonials. For additional case studies and quotes, please visit the following pages on DataDirect's site:

<http://www.xmlconverters.com/customers/use-cases.html>

<http://www.xquery.com/customers/>

Customer Quote: "... We use DataDirect XQuery® as a generic data query tier to separate data access from individual database implementations because we have Oracle, DB2, and MS SQL Server ..."

DataDirect Differentiators, Benefits, and Value Proposition

DataDirect Data Integration Suite delivers multiple benefits:

- **Business flexibility:** One of the top three concerns reported by IT organizations is the need for better business alignment. In the real world, this means the ability to respond flexibly and nimbly to changing business requirements. It also means delivering new products and services quickly, regardless of the challenges of doing so. By simplifying integration of diverse data from multiple sources, DataDirect Data Integration Suite positions IT to develop and implement integration projects faster, improving efficiency as well as business alignment.
- **Better use of expensive technical personnel:** Traditionally, developer productivity has been hampered by the need to write and maintain code, much of it related to data access and manipulation. Often, the same code is repeated multiple times in multiple applications. DataDirect Data Integration Suite addresses this problem by:
 - Minimizing coding by providing a standard mechanism to access heterogeneous data from multiple sources
 - Providing access to data from multiple sources with a single XQuery statement
 - Introducing data access consistency, which reduces the learning curve typically associated with accessing and integrating data from different platforms
 - Enabling developers to focus on logic and business enablement instead of on low level data differences.
- **Solves the problem of integration for both internal and external projects:** Virtually every company engaged in B2B activities supports multiple integrations, both with external companies and between internal systems. While data transformation for use in SOA and Web Services deployments has traditionally required an Enterprise Application Integration or Enterprise Service Bus (ESB) platform, DataDirect Data Integration Suite offers a simpler alternative. Because it is relatively inexpensive and easy to deploy, it is ideal for departmental use, for projects requiring quick turnaround, or for introducing Web Services into the service portfolio.

- **Bi-directional access to virtually any data type:** DataDirect Data Integration Suite streamlines the processes of reading and/or writing diverse data. In conjunction with DataDirect Shadow, this includes mainframe data, which is a must-have capability for many medium to large-sized companies.
- **Lightweight suite, quick deployment:** Deployment is quick and does not require installation of a platform or expensive consulting resources. The Suite is affordable and easy to learn and use.
- **Standards-based:** “Translates” all data formats to XML, providing a straightforward means to combine and manipulate data from dissimilar data files into an industry-standard format.
 - Provides an upgrade path: DataDirect offers a fast, easy way for companies to migrate from legacy systems to newer technologies and architectures.
 - Extends XQuery standard: Data Direct XML Converters and XQuery extensions provide access to multiple data types, a capability that pure XQuery lacks.
 - Performance and reliability: Since XML is text-base and bulky, it can create performance challenges. DataDirect adds stability and reliability and is optimized for scalability and performance.
- **DataDirect history and reputation as a long-standing vendor of high performing data access solutions:** DataDirect has an enviable track record as a vendor offering quality tools for a wide variety of heterogeneous relational databases. Branching out into standards-based XML and other file types benefits customers in virtually all industries. Because DataDirect is an independent supplier, not a suite vendor, the company is highly motivated to provide converters for the widest range of data sources possible.
- **High quality product and customer support:** DataDirect’s extensive training and support resources are cited by customers as significant differentiators
- **ROI from multiple sources:** Multiple customer stories attest to the fact that this suite of products yields multiple benefits. Developer productivity, low acquisition and maintenance costs, and high performance all contribute to cost optimization. In addition, improved business flexibility and faster time to market enable the business to become more competitive and differentiating.
- **Try before you buy:** DataDirect offers a fifteen day trial download at <http://www.xquery.com/download/ddis>.

Customer Quote: “... *Since DataDirect XQuery® doesn’t require WebLogic or any other application server, our customers can decide if they need an application server ...*”

EMA Perspective

Technology heterogeneity translates to complexity, and complexity in turn translates directly to cost. Mitigating these factors with high-quality enterprise management solutions is the way of the future, and a key enabler for IT organizations seeking to take advantage of increasingly sophisticated technology.

The reality is, however, that management products themselves are often expensive, complex, and require months of effort and expensive consulting resources to deploy. Integration solutions are no different, and it is refreshing to see a vendor take a common sense approach to this problem.

Already well-regarded for its Relational technology, DataDirect's expansion to the standards-based world is a boon to IT organizations struggling with heterogeneity. In addition to product functionality and extensive platform coverage, customers also benefit from a highly regarded technical support organization and a company that obviously stays ahead of the changing needs of the industry. This is borne out by multiple testimonials from customers willing to step up to the plate with endorsements, virtually all of which include impressive time, revenue, and/or cost-saving benefits.

Multiple roles within a company will find this product useful. Development organizations are the obvious choice, as many are finding that heterogeneous technology is impacting productivity. Architects will find the Suite equally valuable, as it gives them leeway to effectively utilize data from new sources. Using the capabilities the Suite offers, businesses can embark on a new era of flexibility fueled by faster development turnaround times and new versatility in terms of data utility.

The DataDirect Data Integration Suite does have its limitations. It is not intended to be a Data Integration Platform and lacks platform capabilities such as governance, replication, and process integration. On the other hand, it does fulfill its designated purpose – as a lightweight data integration solution – very well.

Companies seeking an easy route to integration, those venturing into the realms of SOA and Web Services, and those tasked with a list of integration projects they don't have time to carry out are well advised to download and test this Suite. It brings with it many of the benefits of larger, more expensive, platform-based solutions at a fraction of the cost, and is well worth an evaluation.

About DataDirect Technologies

DataDirect Technologies is a leading provider of comprehensive software products for connecting the world's most business critical applications to data and services. Running on any platform, using proven and emerging standards, DataDirect Technologies delivers connectivity to an unparalleled range of data sources using standards-based interfaces such as ODBC, JDBC™ and ADO.NET, XQuery and SOAP. More than 300 leading independent software vendors and thousands of enterprises rely on DataDirect Technologies to simplify and streamline data connectivity for distributed systems and reduce the complexity of mainframe integration. DataDirect Technologies is an operating company of Progress Software Corporation (NASDAQ: PRGS). For more information, visit www.datadirect.com.

About Enterprise Management Associates, Inc.

Founded in 1996, Enterprise Management Associates (EMA) is a leading industry analyst and consulting firm dedicated to the IT management market. The firm provides IT vendors and enterprise IT professionals with objective insight into the real-world business value of long-established and emerging technologies, ranging from security, storage and IT Service Management (ITSM) to the Configuration Management Database (CMDB), virtualization and service-oriented architecture (SOA). Even with its rapid growth, EMA has never lost sight of the client, and continues to offer personalized support and convenient access to its analysts. For more information on the firm's extensive library of IT management research, free online IT Management Solutions Center and IT consulting offerings, visit www.enterprisemanagement.com.

This report in whole or in part may not be duplicated, reproduced, stored in a retrieval system or retransmitted without prior written permission of Enterprise Management Associates, Inc. All opinions and estimates herein constitute our judgement as of this date and are subject to change without notice. Product names mentioned herein may be trademarks and/or registered trademarks of their respective companies. "EMA" and "Enterprise Management Associates" are trademarks of Enterprise Management Associates, Inc. in the United States and other countries.

©2009 Enterprise Management Associates, Inc. All Rights Reserved. EMA™, ENTERPRISE MANAGEMENT ASSOCIATES®, and the mobius symbol are registered trademarks or common-law trademarks of Enterprise Management Associates, Inc.

Corporate Headquarters:
5777 Central Avenue, Suite 105
Boulder, CO 80301
Phone: +1 303.543.9500
Fax: +1 303.543.7687
www.enterprisemanagement.com



1844.030209