

Turning fragmented data into a single view of your customers, suppliers, and products.

## RLP Technologies: Supporting Large XML Files

### Customer Profile

RLP Technologies, Inc., (RLPT) is a wholly owned subsidiary of R.L. Polk & Co., the premier provider of automotive information and marketing solutions.

RLPT's flagship solution, OneView360°, is an automated data management and database production application. This comprehensive data factory software application is designed to automate the collection of data from disparate sources, create standardized data feeds, orchestrate the enrichment of data to improve its accuracy, create a "Single Source of Truth" for all of an organization's information assets, and distribute the data for use in analytical and operational applications. The solution is both an engine that enables the timely processing of data, and a single integrated application that provides IT, data and business analysts one user interface in which to work to manage their data.

Some of the key benefits attained from implementing OneView360° are to provide the most stable, predictable, flexible and cost-effective data production environment. With an open, service-oriented architecture (SOA), OneView360° can seamlessly integrate investments in legacy, commercial off-the-shelf applications and web services.

### Understanding the Business Challenge

RLPT's goal is to provide their customers a competitive advantage through accurate and timely business information. The ability to quickly analyze and leverage all facets of data within the business and make the appropriate decision faster than your competitors can put you ahead of the game.

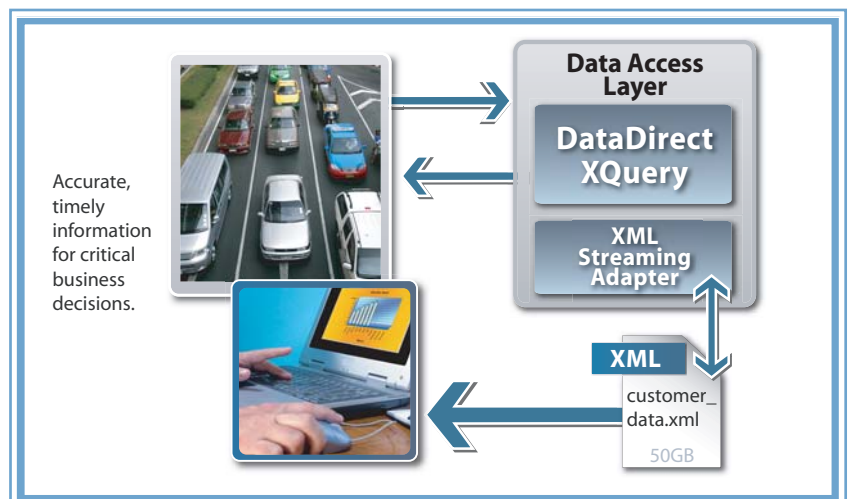
Information for these business decisions is only as good as the data. The quality of the data and the ability to process it quickly and cost-effectively are important factors driving the information supplied to customers.

OneView360° is a data processing engine capable of moving very large volumes of data around in XML. For example, in one implementation of OneView360°, source data files containing automotive industry customer records can exceed 30GB, and can exceed 50GB after the source data is enriched.

A Data Quality Inspection module is a critical component of OneView360°, which incorporates checkpoints throughout the entire application. The early identification of data anomalies and quality issues is a critical success factor. With DataDirect XQuery<sup>TM</sup>, OneView360° is able to perform data profiling functions such as complex frequency distributions, counts and pattern analysis data directly on the large XML data files.

"DataDirect's optimization techniques including Document Projection and Document Streaming allowed us to meet our requirements and clearly positioned DataDirect XQuery ahead of its competitors."

Clara Sagan,  
Senior Product Engineer,  
RLP Technologies



## Choosing the Right Solution – DataDirect XQuery<sup>™</sup>

A major requirement of OneView360<sup>°</sup> was to support the XML documents in their native format and to support them regardless of size. Other solutions require intermediate processing of the XML data files. This can add to the time required for both development and processing, with the potential for scalability issues.



RLPT found that the XQuery approach to handling and processing XML was one of the best approaches to dealing with the extraction and manipulation of information in native XML format. XQuery provided the means for collecting statistics about the XML content by enabling sophisticated querying capabilities which allowed OneView360<sup>°</sup> to evaluate measurements taken on the file against history and preset quality thresholds.

The issue of large XML file support caused RLPT to test and analyze various solutions. Ultimately, DataDirect XQuery provided the fastest and most scalable XQuery implementation for RLPT's current requirements and projected future growth. During testing, DataDirect XQuery processed data at least seven times faster than other solutions. Other XQuery products showed performance degradation and high memory consumption when processing much smaller XML files.

"To meet the product requirements for OneView360<sup>°</sup> we needed to process XML files of up to 50GB successfully 100% of the time in an efficient manner. DataDirect's optimization techniques including Document Projection and Document Streaming allowed us to meet our requirements and clearly positioned DataDirect XQuery ahead of its competitors," said Clara Sagan, senior product engineer, RLPT.

## Realizing the Benefits of DataDirect XQuery<sup>™</sup>

"DataDirect XQuery provided us a commercial, well-supported implementation of the XQuery specification while delivering the performance, scalability and reliability necessary for our large scale XML profiling needs," said Joe LaFeir, vice president of product development, RLPTechnologies. "Using the DataDirect product has helped us significantly reduce our development effort, which helped bring our OneView360<sup>°</sup> Data Quality Inspection module to market faster."

After testing alternative solutions and products, RLPT chose DataDirect XQuery for these reasons:

- DataDirect XQuery processed the XML documents with outstanding performance. OneView360<sup>°</sup> is able to process the XML documents directly, which is the quickest way to deliver business-critical information.
- DataDirect XQuery was a proven, reliable XQuery product. Other products and solutions couldn't meet RLPT's XML file size requirements. DataDirect XQuery's optimizations of memory usage and consumption gave OneView360<sup>°</sup> the ability to meet and exceed required file sizes.
- RLPT developers found DataDirect's technical support to be a valuable asset. The technical support team worked closely with RLPT to provide best practice strategies when working with large XML files and quickly resolved their questions and issues in a timely manner.

"Using the DataDirect product has helped us significantly reduce our development effort, which helped bring our OneView360<sup>°</sup> Data Quality Inspection module to market faster."

Joe LaFeir, Vice President of Product Development, RLPTechnologies.

DataDirect XQuery enabled RLPTechnologies to process very large XML data files natively to efficiently measure the quality of data, while allowing them to reduce development time.

### About DataDirect Technologies

DataDirect Technologies is the software industry's only comprehensive provider of software for connecting the world's most critical business applications to data and services, running on any platform, using proven and emerging standards. Developers worldwide depend on DataDirect<sup>®</sup> products to connect their applications 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 to 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](http://www.datadirect.com).

DataDirect and DataDirect XQuery are trademarks or registered trademarks of DataDirect Technologies Corp. in the U.S. and other countries. Java and all Java based marks are trademarks or registered trademarks of Sun Microsystems, Inc. in the U.S. and other countries. Any other trademarks contained herein are the property of their respective owners.