(intel®)

Next-Gen Data Platform Powering Healthcare Modernization and Innovation

Drive better, faster, and less expensive data integration and infrastructure modernization projects through the MarkLogic* Enterprise NoSQL database platform running on Intel® architecture

This solution brief describes how to solve business challenges and enable digital transformation through investment in innovative technologies.

If you are responsible for...

- **Business strategy:**
  You will better understand how a NoSQL database platform solution will enable you to successfully meet your business outcomes.

- **Technology decisions:**
  You will learn how a NoSQL database platform solution works to deliver IT and business value.

Executive Summary

Healthcare organizations today are striving to stay relevant by improving customer satisfaction, reducing costs, delivering better outcomes for patients, and building a modern sustainable enterprise. These business imperatives have increased in importance since the passage of the Affordable Care Act (ACA). The recent implementation of the Medicare Access and CHIP Reauthorization Act (MACRA), with their focus on driving value-based care, has put additional pressure on healthcare organizations—especially payers, but also providers—to adapt and stay competitive. MarkLogic has been cited by multiple industry analysts as a leader in the operational and NoSQL database markets.[1]

However, dealing with the sheer volume and diversity of standalone, proprietary healthcare systems and heterogeneous data is a significant hurdle to achieving healthcare business objectives. Data is everywhere. Patient-generated data, claims and billing data, clinical data, reference source data, and more are stored in an increasing variety of systems, devices, and locations. Aggregating, harmonizing, and enriching data to get a 360-degree view of a member or patient is more crucial, yet more difficult, than ever. In addition, as the volume and velocity of data continue to grow, the challenge to healthcare organizations only increases.

The MarkLogic* Enterprise NoSQL database platform running on the latest generation of Intel® Xeon® processors creates an agile, modern, scalable platform for innovation across the healthcare enterprise.
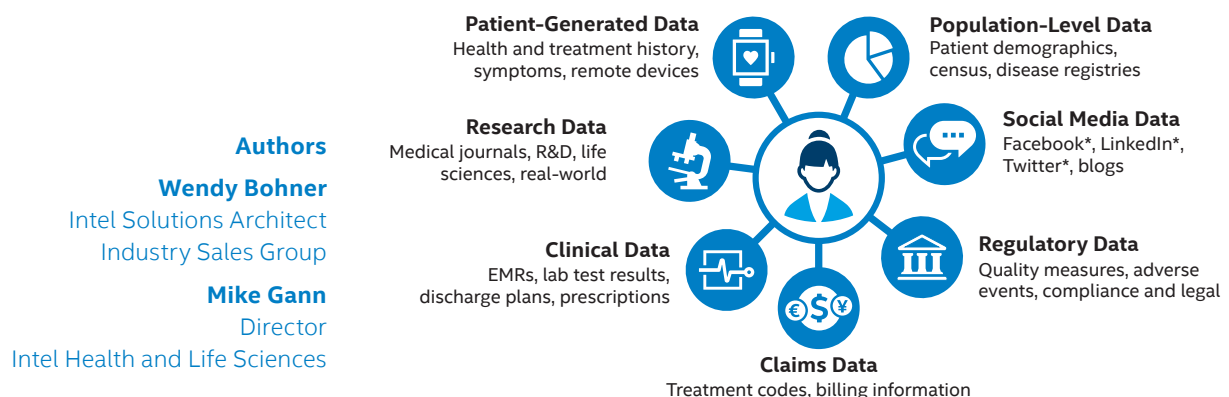
Authors

Wendy Bohner
Intel Solutions Architect
Industry Sales Group

Mike Gann
Director
Intel Health and Life Sciences

MarkLogic®

Patient-Generated Data
Health and treatment history, symptoms, remote devices

Population-Level Data
Patient demographics, census, disease registries

Research Data
Medical journals, R&D, life sciences, real-world

Social Media Data
Facebook*, LinkedIn*, Twitter*, blogs

Clinical Data
EMRs, lab test results, discharge plans, prescriptions

Regulatory Data
Quality measures, adverse events, compliance and legal

Claims Data
Treatment codes, billing information

Figure 1. Healthcare payers need a secure platform that can integrate a wide variety of data no matter where it lives or whether it is structured or unstructured.

Solution Benefits

- Quickly integrate data across silos and domains.
- Get answers faster.
- Improve time to value and competitiveness.
- Improve compliance, security, and privacy.

Business Challenge: Integrating Data and Getting Answers Quickly

Healthcare data is growing at an almost unimaginable pace, outstripping most other industries with its 48-percent annual growth rate. By 2020, there may be more than 2,300 exabytes of healthcare data worldwide. What's more, by 2020 over 50 percent of electronic medical record (EMR) systems are expected to be replaced by next-generation systems that meet the ever-changing compliance demands and data formats.[2] Unstructured data, such as audio, video, and images, are increasingly stored within medical records, becoming key parts of patient records.

The various sources and types of data, such as patient-generated data, research data, clinical data, claims data, demographics, social media data, reference data, and regulatory data, often sprawl across disconnected relational database management system (RDBMS) silos. Without a comprehensive view across the enterprise, it is difficult to use the data to achieve business goals, operate efficiently, and meet the increased demand for better service and value from consumers and business partners. As data sources and analytics requirements evolve, it can be time-consuming and costly to change traditional RDBMS schema. Data scientists, according to interviews and expert estimates, spend from 50 to 80 percent of their time just collecting and preparing unruly digital data, before it can be explored for useful nuggets.[3]

In addition, increased competition, evolving regulatory requirements, and new demands from payers, providers, and consumers have contributed to a growing trend of consolidation across the healthcare industry, driven by a desire to gain economies of scale. As the healthcare consolidation trend continues, data integration challenges become even greater as companies struggle to coalesce disconnected legacy IT systems. It is not unusual for a data integration project to take five years, and many are never completed. These delays prevent an organization from gaining a competitive edge in today's fast-paced, crowded healthcare marketplace.

Simply put, healthcare organizations need a flexible platform that can keep up with the industry's data velocity and changing business models.

Using Data to Speed Innovation and Empower Decision Making

Healthcare payers are already cutting costs, saving time, and gaining competitive advantages with the MarkLogic* Enterprise NoSQL database platform. MarkLogic has been cited by multiple industry analysts as a leader in the operational and NoSQL database markets.[4] ZDNet recently published an article that stated, "The notion of having one database to process both transactional and analytical workloads has been a recurring one through the years… MarkLogic is dual-wielding its way through transactional and analytical workloads…accommodating all kinds of data, and it's flexible enough to make the need to know your schema in advance obsolete."[5]

According to *Upside*, published by The Data Warehousing Institute (TDWI), extract, transfer, and load (ETL) software and processes may consume up to 80 percent of the total cost of a data warehouse project.[6] By avoiding the extensive up-front data modeling and ETL processing typical of RDBMS data, payer organizations can get data into—and answers from—the systems faster. Innovation starts with leveraging all of an organization's data without the burden of ETL, because an organization can ingest a large volume of relational data from different repositories with various schemas. With MarkLogic, there is no need to prebuild a master data model or identify all relevant queries beforehand.

Advanced full-text search-and-query with semantics capabilities let payers leverage data from across an organization in new and innovative ways in record time. For example, payers can use the MarkLogic solution to quickly build applications that empower better decision making. Healthcare providers that offer their own health insurance plans seek opportunities for early patient interventions to improve health outcomes. Using MarkLogic's Enterprise NoSQL database platform, these providers can build holistic and semantically rich views of their patients. They accomplish this by linking and contextualizing relationships between dozens of medical and nonmedical data sources, including clinical, EMR, behavioral, prescriptions, demographics, claims, and contact history. They can also analyze EMRs alongside key data not traditionally stored in EMRs, such as social determinants data, in real time.

By building tailored tools that leverage data in a centralized clinical data repository, healthcare organizations can proactively implement new programs that improve the health of their customers. Other use cases for the MarkLogic Enterprise NoSQL database platform include the following: data integration; mainframe migration; consumer engagement and retention; client management; network management; payment and reimbursement; population health and medical management; and quality, accreditation, and compliance.

GET ANSWERS FASTER

2X

The MarkLogic* Enterprise NoSQL database, running on high-performance Intel® Xeon® processors and Intel® Solid State Drives, **speeds query performance by more than 2x from previous generations of Intel Xeon processors**.[7]

Solution Value: Efficiency and Cost Savings

As the leading next-gen data technology and Enterprise NoSQL database, MarkLogic's solution has flexible data models, so payer organizations can load their data into the data platform as-is without resource-intensive ETL. Because the MarkLogic NoSQL database platform indexes the data as it is loaded, payers can get meaningful answers from their data in real time through the platform's built-in search and query capabilities. Ideas go to market faster.

The MarkLogic Enterprise NoSQL database platform alleviates the pain of integrating healthcare data. As an agile data integration platform, MarkLogic's solution enables healthcare organizations to better utilize all of their data to quickly create a 360-degree view of a health consumer. Compared to the traditional ETL-focused RDBMS, MarkLogic's Enterprise NoSQL database platform offers the following benefits:[8]

- **Reduces complexity.** Flexible data models enable organizations to integrate unstructured and structured data without the time and expense of up-front data modeling. Organizations can easily integrate data from an existing RDBMS with data in other common formats, such as JSON, XML, RDF, geospatial, text, video, images, and more. This flexibility enables healthcare organizations to adapt to rapid changes in data volume and velocity.

- **Provides answers more quickly.** Sophisticated indexing capabilities support full-text search and query— immediately upon ingestion. Users can search across all their data quickly, using JavaScript*, XQuery*, and SPARQL* directly within MarkLogic's NoSQL database. It is easy to access both structured and unstructured data in real time.

- **Empowers better decisions.** Built-in powerful semantics help identify relationships across all of an organization's data. This enables organizations to connect disparate data to infer new facts without modifying the source data.

- **Improves competitiveness and efficiency.** Efficient data processing provides a competitive edge. Healthcare organizations achieve a quicker return on investment, complete data integration projects faster, and spend more time on innovation and less time on manual data processes that slow down key business processes. With MarkLogic, projects show results more quickly, scale easier, and can change and adapt as needed with lower costs and quicker turnaround.[9]

MarkLogic* Enterprise NoSQL Database Platform Eases Integration Pains, Improves Business Efficiency

Following a merger, a top-five commercial payer needed to integrate more than 200 different data sources for human resources (HR) data into a central location. Facing a five-year integration project, the company turned to MarkLogic. By implementing MarkLogic's Enterprise NoSQL database platform, the company was able to centralize data from those sources and deliver it to more than 50 downstream systems in only a year. The resulting data platform handles terabytes of HR data daily and enables maximum reuse and consistency of data throughout the organization.[10]

For another leading healthcare payer, MarkLogic's Enterprise NoSQL database platform provided rapid data integration across multiple data silos to create a central document repository of over 300,000 communication sources—calls, chats, texts, mail, fax, and email. Customer service representatives now have a 360-degree view and can access and search this repository in real time during member engagements, enabling them to obtain a comprehensive view of inbound communications at sub-second speeds. It took only four months to develop this capability using MarkLogic's solution.

With the MarkLogic NoSQL database platform, healthcare organizations do not need to rely on error-prone manual processes to drive client onboarding. By linking and contextualizing data and metadata with semantics, the solution can automate and streamline data discovery and workflow processes, dramatically reducing the need for manual data entry. Several organizations focused on client satisfaction have built their recommendation engines using the MarkLogic database platform's embedded semantics features.[11]

MarkLogic Enterprise NoSQL Database Platform

Relational Databases

Mainframes

File Systems

Any Other Source

MarkLogic®

Transactional Applications

Analytical Applications

Downstream Systems

- **Keeps data more secure.** The MarkLogic Enterprise NoSQL database is the only NoSQL platform to earn a Common Criteria security certification. Fine-grained, role-based security at the document level helps protect data against threats and has a proven track record of running mission-critical applications in secure environments. The platform is certified to run in classified government systems.

MarkLogic's Enterprise NoSQL database platform accelerates access to healthcare data wherever that data is located. When running on the latest Intel® Xeon® processors, an organization can get answers from its data even faster. Benchmarks demonstrate that running MarkLogic's Enterprise NoSQL database on servers equipped with the Intel® Xeon® processor E5-2699 v4 and using the Intel® Solid State Drive (Intel® SSD) Data Center P3700 Series for storage increases query performance more than 30 percent compared to using the Intel® Xeon® processor E5-2697 v3. Queries are more than twice as fast as using an even older previous-generation Intel Xeon processor.[12,13]

Solution Architecture: Quick, Affordable, and Secure Data Integration

MarkLogic's Enterprise NoSQL database platform lies at the center of a solution stack that includes a top layer of applications and data delivery tools, and supporting layers featuring Apache Hadoop* open source software optimized for Intel® architecture (see Figure 2). Here are some details about each layer of the solution stack, starting from the bottom:

- **Intel architecture.** High-performance yet affordable Intel Xeon processors and reliable, fast Intel SSDs provide the compute and storage capabilities that power the data management platform. Intel architecture supports the speed necessary for real-time queries, as well as the number crunching brawn required for large batch workloads.

- **Hadoop.** For organizations that have made an investment in Hadoop*, MarkLogic can complement existing Hadoop technology to improve its governance, security, and operational value. Designed precisely for massively distributed storage and the processing of very large data sets and optimized for the latest generation of Intel® processors, Hadoop is inexpensive, highly scalable, and can handle raw, prepared, and indexed data. Hadoop's affordable, scalable data lake eliminates the need for ETL. MarkLogic enhances the value of the data in Hadoop by making it more secure, accessible, and operational. While MarkLogic deployments typically focus on real-time data access and analytics, rather than staging, ETL, and batch, appropriate Hadoop use can augment MarkLogic in these areas, reducing TCO and increasing agility. For example, older, less-valuable data can be moved from MarkLogic to a separate Hadoop cluster for slower batch processing, while retaining the full power of MarkLogic for an organization's most valuable data.

- **Data management platform.** The MarkLogic Enterprise NoSQL database platform is built to handle search, query, and analytics of native-state, multi-source data and can easily ingest new data sources. The platform can complement an existing Hadoop ecosystem; it runs on affordable, commonly available servers, and is highly scalable—healthcare organizations can add and remove data and processing nodes as needed (one MarkLogic customer has over 280,000 concurrent users). And because of its distributed design, the platform is highly reliable and supports cluster failover to eliminate single points of failure. Role-based security at the document level helps keep data protected against threats.

- **Applications and data delivery.** The combined solution creates a flexible platform to deliver line-of-business applications that can take advantage of the entire healthcare data landscape, including unstructured data that otherwise would be inaccessible for analytics.
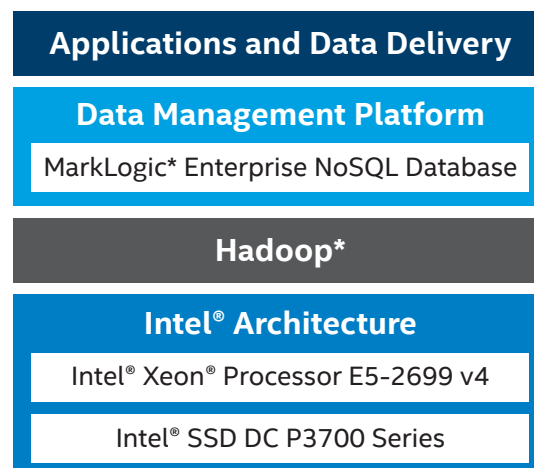
Applications and Data Delivery

Data Management Platform
MarkLogic* Enterprise NoSQL Database

Hadoop*

Intel® Architecture
Intel® Xeon® Processor E5-2699 v4
Intel® SSD DC P3700 Series

Figure 2. The solution is built with affordable, high-performance Intel® hardware, open source Apache Hadoop* distributed storage and processing, and the MarkLogic* Enterprise NoSQL database platform. These layers combine to deliver line-of-business applications that provide accelerated and cost-effective business insight based on both structured and unstructured data from many sources.

Conclusion

To overcome the challenges facing the healthcare industry today, payers must find ways to improve performance, increase value, and reduce risk. With competitors large and small seeking to gain a competitive edge, it is critical to realize faster time-to-value from healthcare data spread across the organization. The MarkLogic NoSQL database platform empowers healthcare organizations—payer, provider, and health IT vendors—to quickly ingest, index, search, enrich, and extract value from their structured and unstructured data across silos and domains.

A crowded marketplace and unprecedented consolidation activity in the healthcare industry present healthcare organizations with challenges beyond getting value out of their own data silos. To realize the value of scaling up the business, these organizations now must find ways to integrate into a single infrastructure multiple systems performing the same functions. The shorter the integration timeline, the quicker the organization can realize the business value of the deal. Data integration, and the business intelligence that comes with it, is often the key metric of the success of a consolidation.

MarkLogic's Enterprise NoSQL database platform, running on high-performance Intel Xeon processors, is designed to meet today's healthcare business challenges. It quickly extracts value from data and provides a more secure platform for leveraging a payer organization's healthcare data, no matter where that data lives and whether it is structured or unstructured.

Find the solution that is right for your organization. Contact your Intel representative or visit **intel.com/ healthcare.**

Solution Provided By:

Focus on MarkLogic

MarkLogic® is the world's best database for integrating data from silos. MarkLogic is an operational and transactional Enterprise NoSQL database platform, and organizations around the world rely on it to integrate their most critical data and build innovative applications on a 360-degree view of that data. MarkLogic makes it easy to get data in and easy to get data out, and it is 100 percent trusted to run critical business operations. Headquartered in Silicon Valley, MarkLogic has offices throughout the U.S., Europe, Asia, and Australia.

MarkLogic helps make the transition to its Enterprise NoSQL database platform successful by offering training, consulting services, expert support, and a vibrant online community.

For more information, please visit marklogic.com.

Solutions Proven By Your Peers

Intel Solutions Architects are technology experts who work with the world's largest and most successful companies to design business solutions that solve pressing business challenges. These solutions are based on real-world experience gathered from customers who have successfully tested, piloted, and/ or deployed these solutions in specific business use cases. Solutions architects and technology experts for this solution brief are listed on the front cover.

Learn More

You may also find the following resources useful:

- Real-Time Performance for Enterprise Applications Solution Brief
- Healthcare Receives Big Data Solutions with MarkLogic and Intel White Paper

[1,4] Gartner.com (October 5, 2016).

[2] "The Digital Universe: Driving Data Growth in Healthcare." EMC and IDC, 2014.

[3] Steve Lohr. "For Big-Data Scientists, 'Janitor Work' Is Key Hurdle to Insights." The New York Times. August 17, 2014.

[5] Zd.net (August 31, 2016).

[6] Adam Wilson. "Let the People Who Know the Data Best Do the Wrangling." Upside, published by The Data Warehousing Institute (TWDI), 2016.

[7,8,12] Performance on the MarkLogic* CPoX (Content Processing over XML) benchmark measured by Intel, February 8, 2016. Tests measured query performance for a large set of Wikipedia* documents in multiple languages using a cluster of three two-socket server nodes with three different server platforms: (1) a baseline server configured with 2 x Intel® Xeon® processor E5-2697 v2 (12 cores, 30 MB cache, 2.7 GHz), result = 1860 queries per second; (2) a previous-generation server configured with 2 x Intel® Xeon® processor E5-2697 v3 (14 cores, 35 MB cache, 2.6 GHz), result = 3109 queries per second; (3) a new server configured with 2 x Intel® Xeon® processor E5-2699 v4 (22 cores, 55 MB cache, 2.2 GHz), result = 4113 queries per second (1.32x faster than the previous-generation server, 2.21x faster than the baseline server). All servers were configured with 128 GB of memory, 8 x hard-disk drives (1 TB, 7200 RPM), seven for data, one for the operating system, 1 x Intel® Solid State Drive Data Center P3700 Series, 10 GbE Network, MarkLogic 8.0-3.2, CentOS* 6.6. Performance estimates of 1.13x for the Intel® Xeon® processor E5-2697 v4 versus the previous-generation server were based on the assumption that per-core performance would be the same on the Intel Xeon processor E5-2697 v4 as on the Intel Xeon processor E5-2699 v4.

[9] "Understanding Total Cost of Ownership," MarkLogic video, April 27, 2015.

[10] "Enterprise Data Strategy in the Healthcare Landscape," MarkLogic white paper, 2016.

[11] "Accelerate Healthcare Client Onboarding," MarkLogic paper, 2016.

[13] "Real-Time Performance for Enterprise Applications," Intel solution brief, 2016.