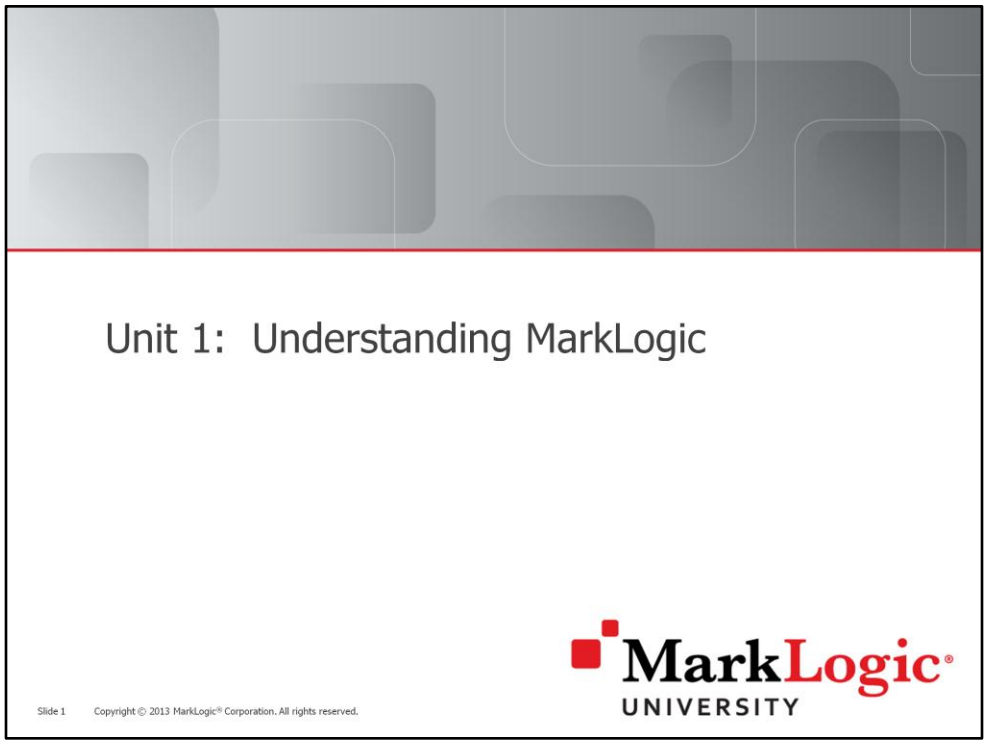
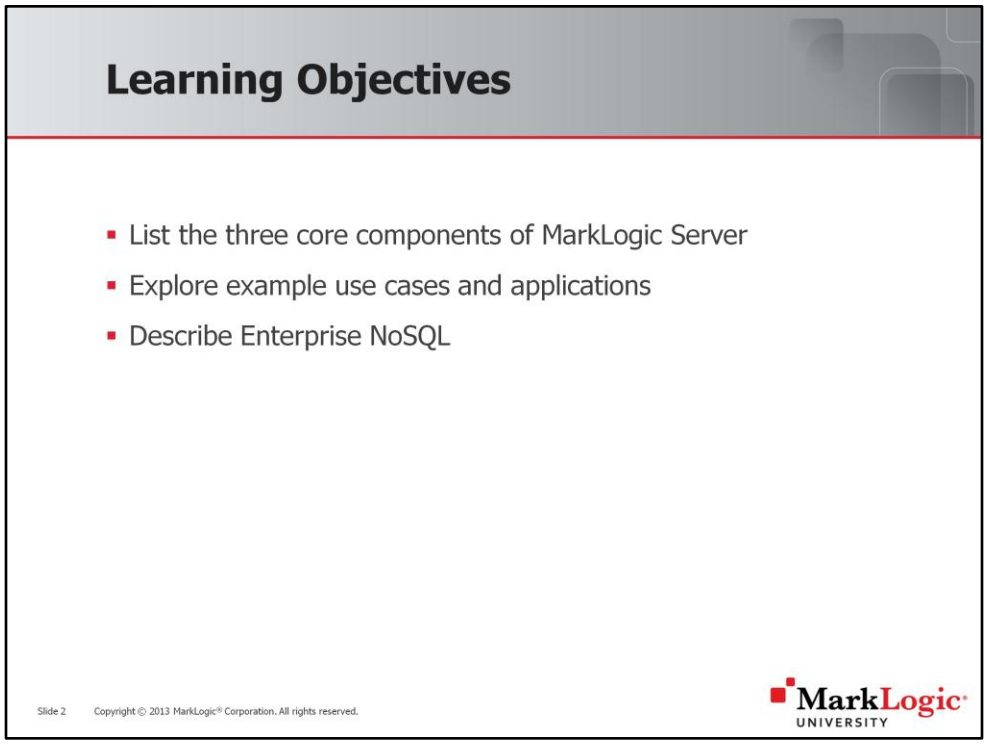
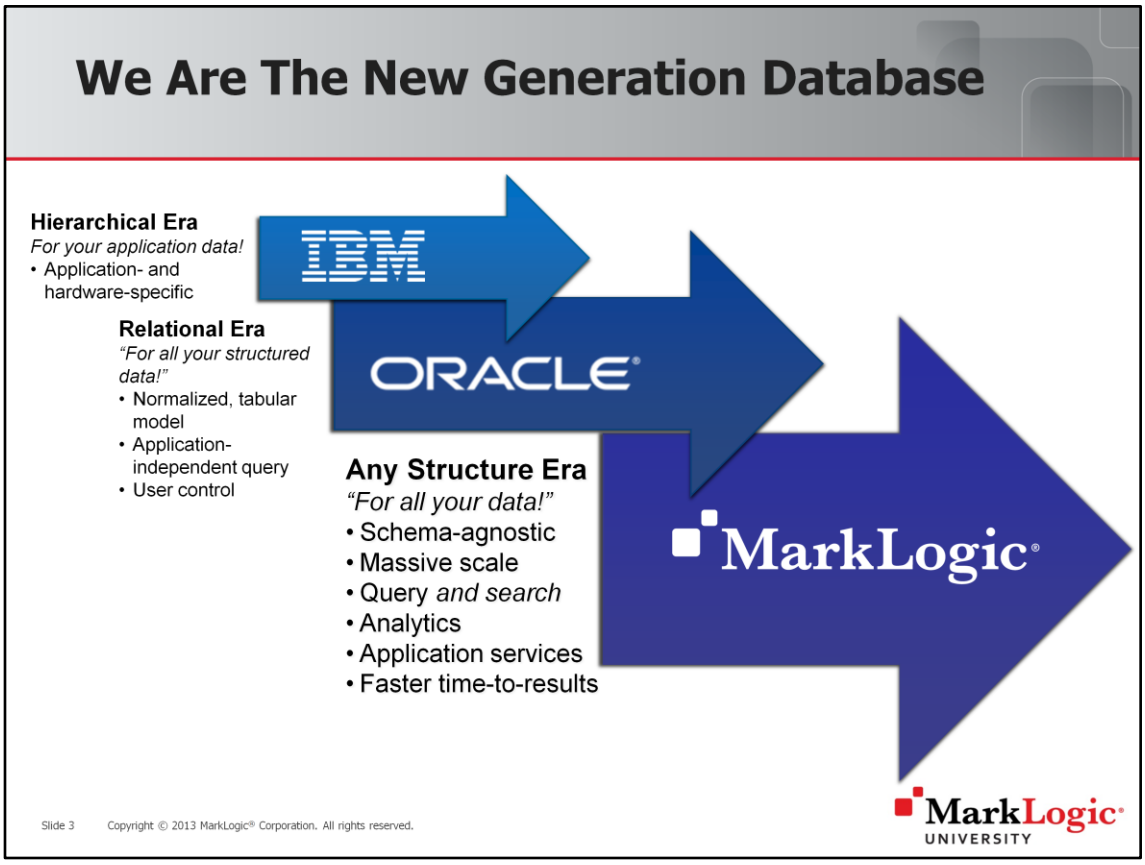
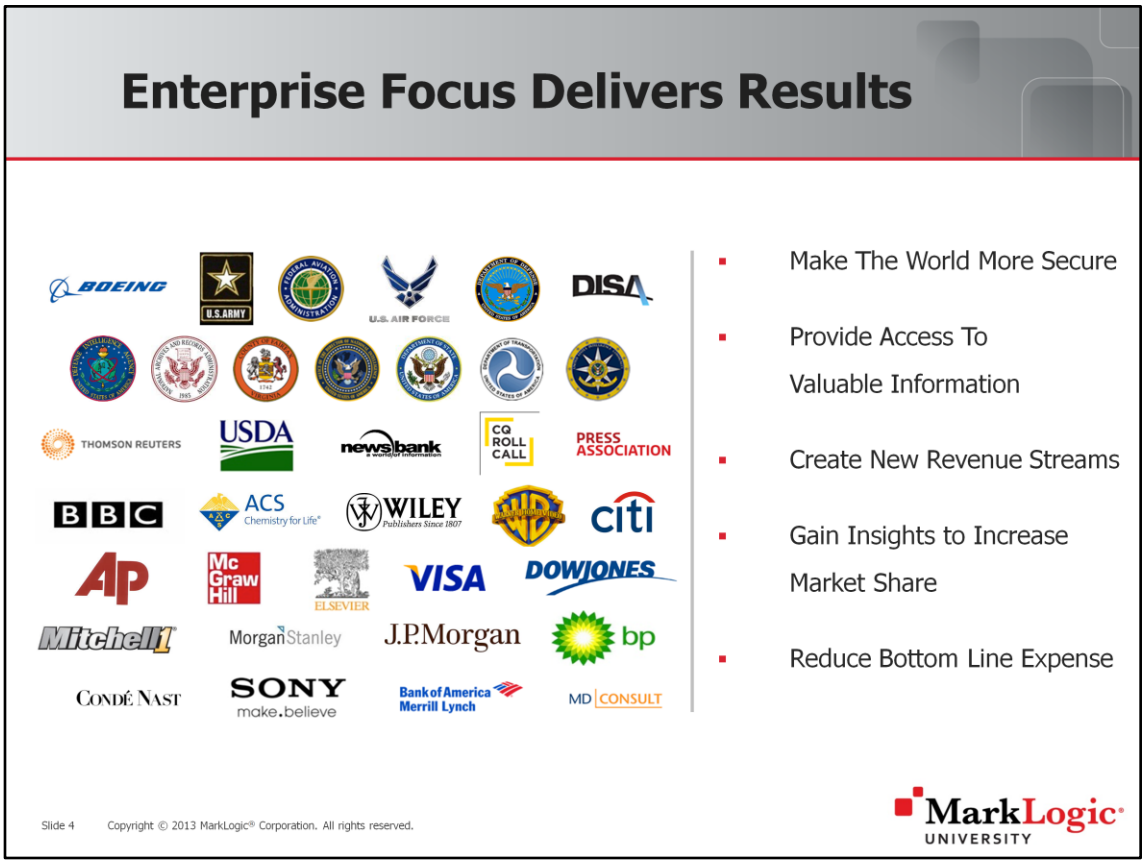






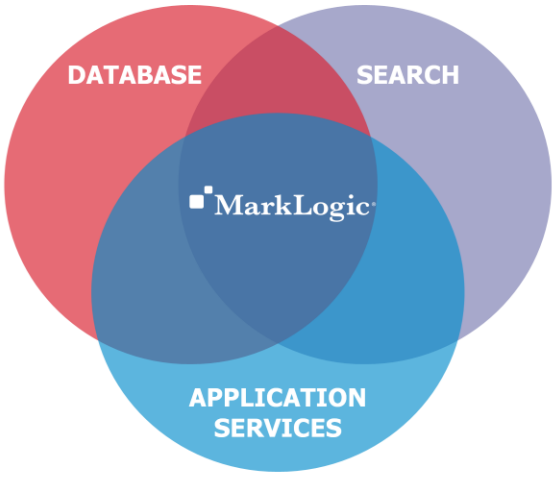


As you look across the MarkLogic customer portfolio, one thing is clear. These are organizations and enterprises that took an early and innovative approach to their big data challenges, and have been able to realize true value.

Some have created new revenue streams by turning existing assets into new products and services.
Some use their data to gain competitive advantage and increase market share by observing and learning about their customers.
Many have been able to reduce their expense line – by consolidating or decommissioning systems – or just by making their teams more efficient.
Our government and intelligence customers have built applications to ensure that we remain safe and secure.
And, finally, you can't ignore the value that you get just by being able to access information that would otherwise be beyond the reach of the people who need it most.

The Only Enterprise NoSQL Database

- Search & Query
- ACID Transactions
- High Availability / Disaster Recovery
- Replication
- Government-grade Security
- Elasticity
- Cloud Deployment
- Hadoop for Storage & Compute
- Semantics



Slide 5

When organizations are looking for infrastructure to manage and leverage Big Data, they look for three things:
A database that can handle unstructured and multi-structured data with ease.

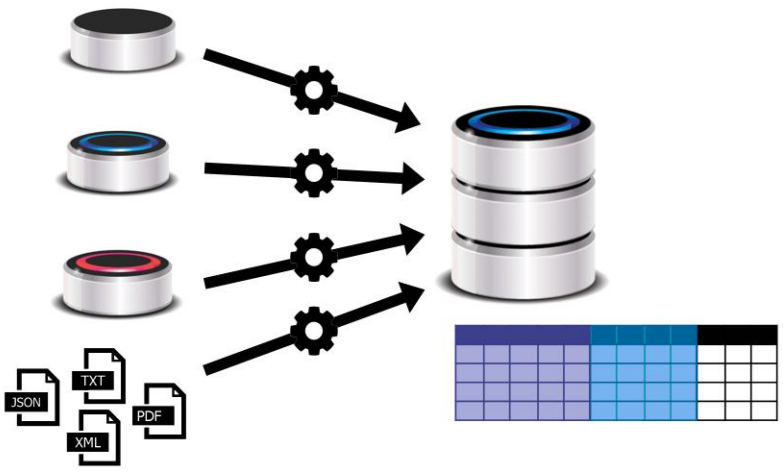
Great search capabilities so users can find the data they are looking for and leverage it to make better decisions for the business.

Application services and tools that allows developers to build applications quickly and easily so that the data turns into usable information.

There are plenty of best of breed technologies out there to serve each one of these functions – but cobbling together a system to do that is time and resource intensive – not only to build, but more so to maintain.
MarkLogic provides all three of those capabilities. And, we have the added bonus of having 11 years under our belt to ensure that the system is enterprise hardened with the security, back up, recovery, high availability and data integrity you come to expect from an Enterprise data management system.

Example Use Cases

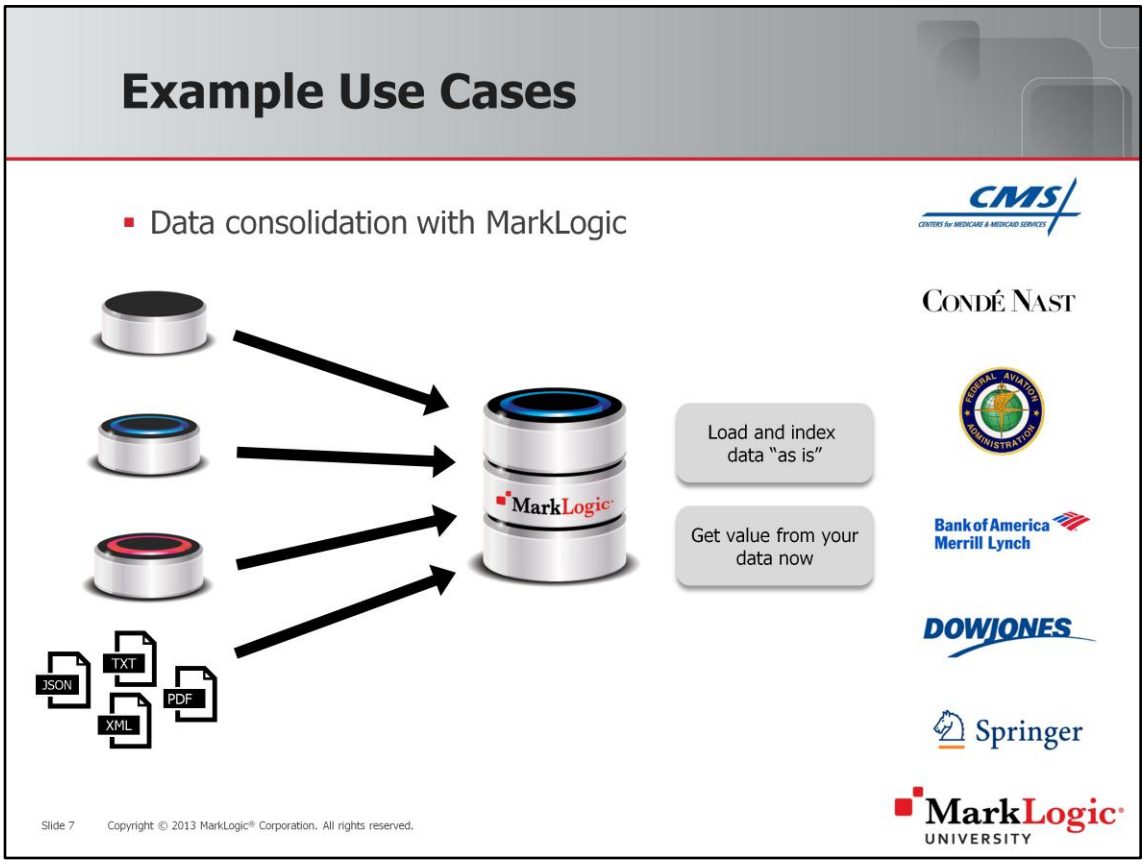
- Data consolidation (before MarkLogic)



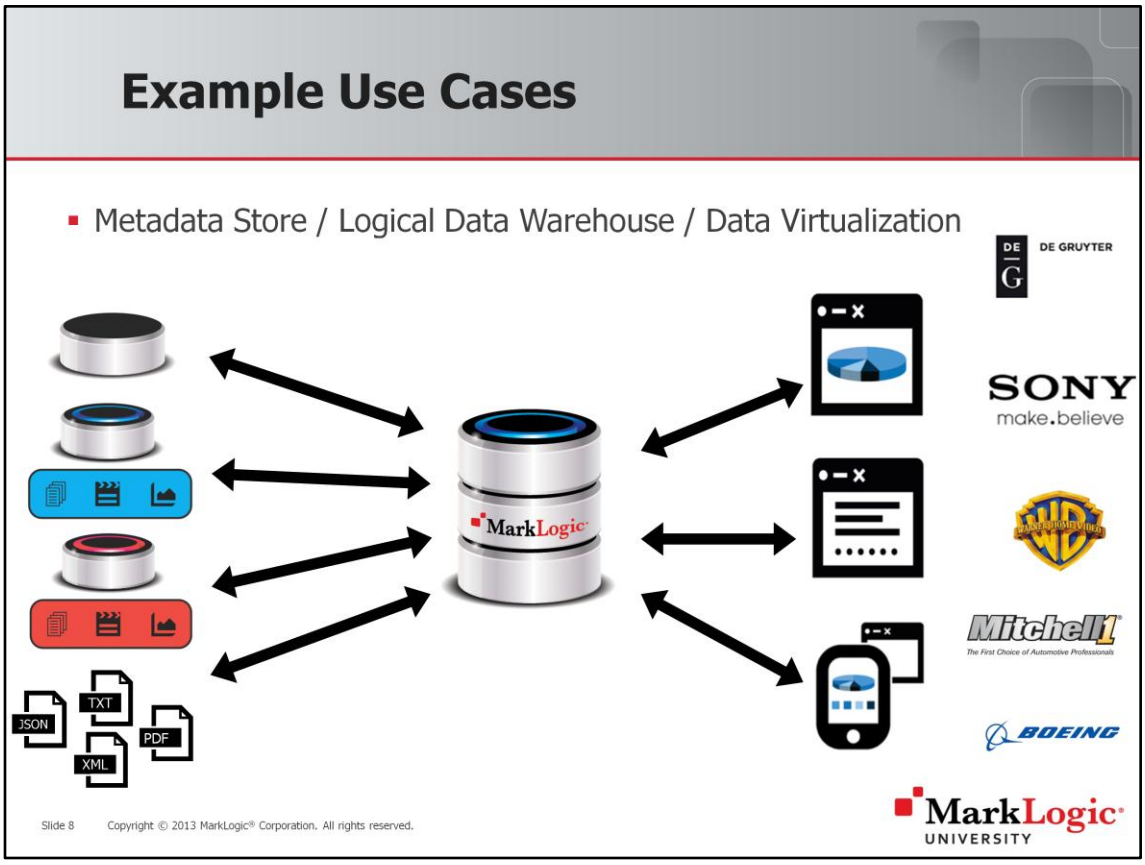
Slide 6

MarkLogic UNIVERSITY

- What people often have today is lots of different siloed systems.
- This is ineffective and costly – a bad mix
- What they'd like to do is consolidate the data into one place
- But then they realize they need a schema that covers all of these disparate systems
- But none of those systems have schemas that are anything like one another, so they have to transform everything
- Then of course they add another data source of one of the schemas change and they have to redefine the transforms and update the schemas
- It's a mess
- With MarkLogic we have a totally different approach



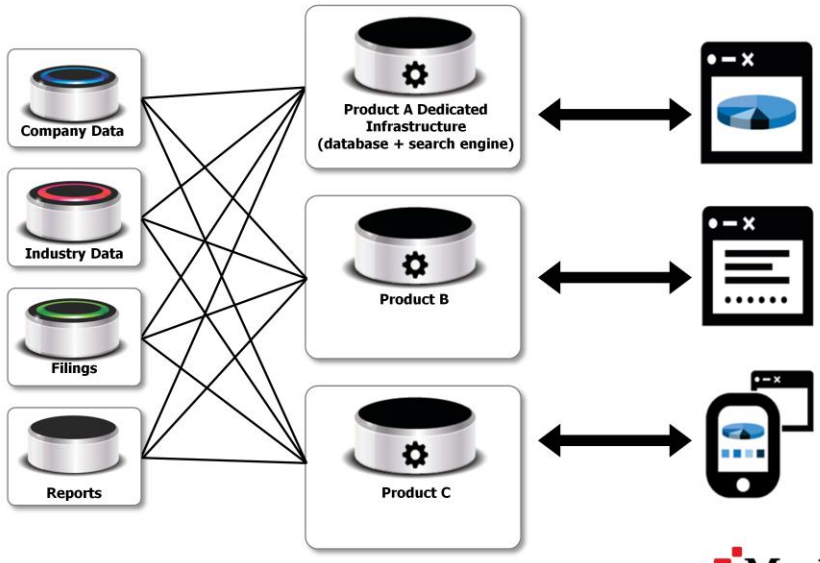
- Because we are Schema-agnostic, we let you load data from all of these different places "as is"
- Don't spend months doing a data modeling exercise up-front.
- Get the data loaded and start using it now
- Here are a few of our customers that use MarkLogic for Data consolidation today



- In some cases you can't or don't want to move all of your data to one place
- For example, you might have some legacy applications running on your old repository and you may need to leave them there
- In this case our customers use MarkLogic as a Metadata Store or Logical Data warehouse
- We pull the metadata from the repositories and let applications access MarkLogic to understand and get value from the underlying resources
- Again, here are a few of our customers that use this pattern
- Sometimes it makes sense to combine these.
    - For example, one customer started off as a Logical Data Warehouse and over time was able to consolidate some of the information from the underlying systems
    - At that point they could turn those systems off which saved them time, complexity, and money

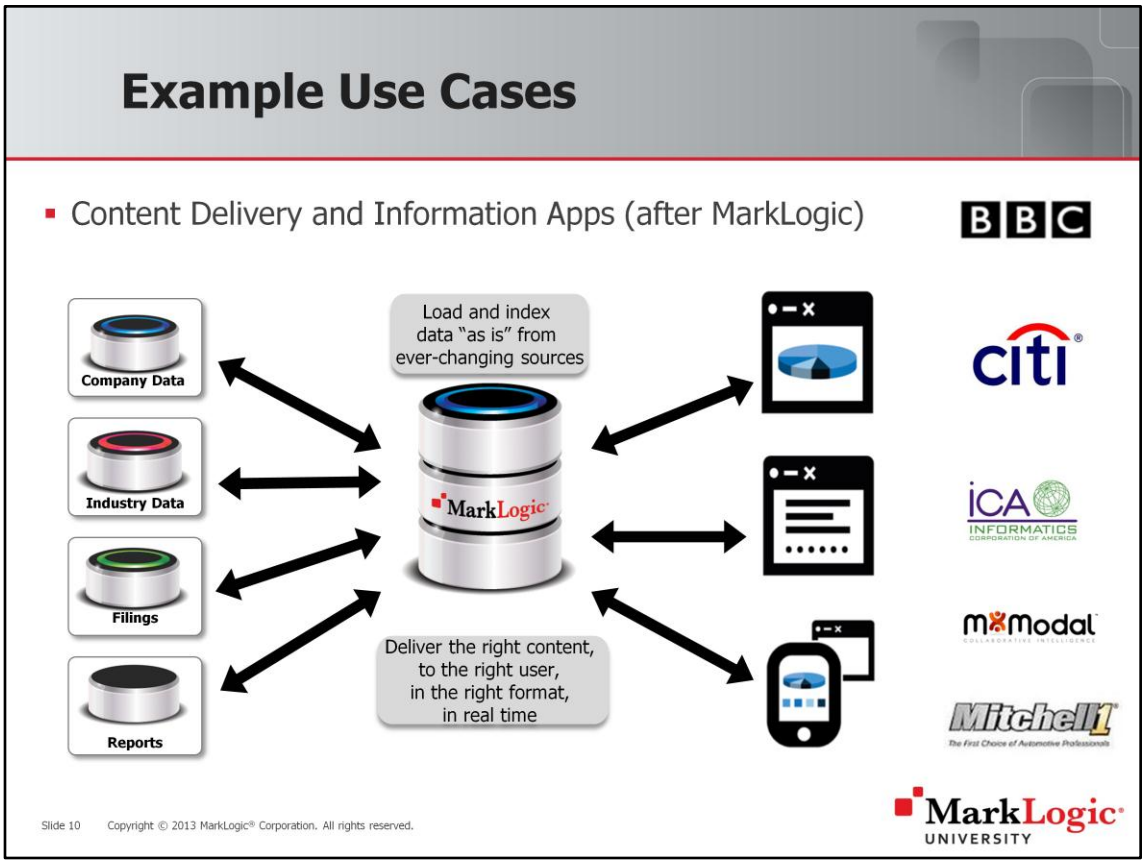
Example Use Cases

- Content Delivery and Information Apps (before MarkLogic)

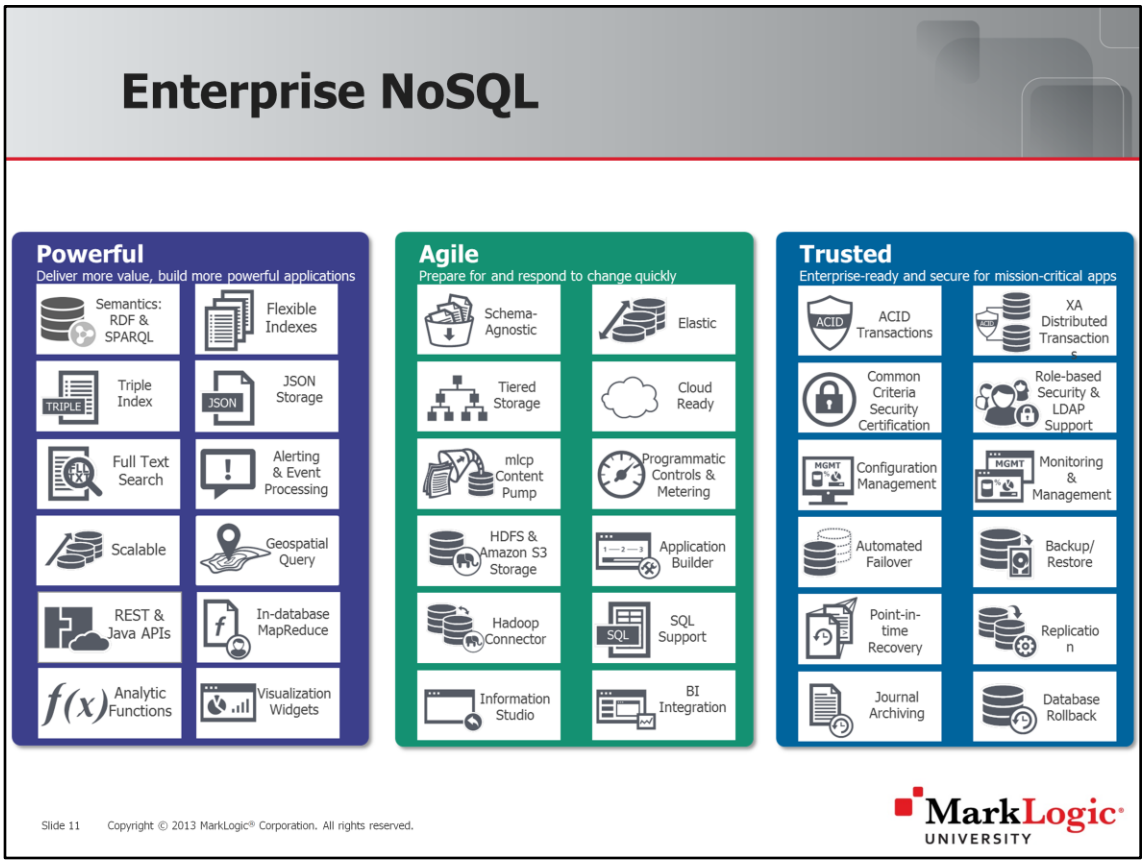


Slide 9
MarkLogic UNIVERSITY

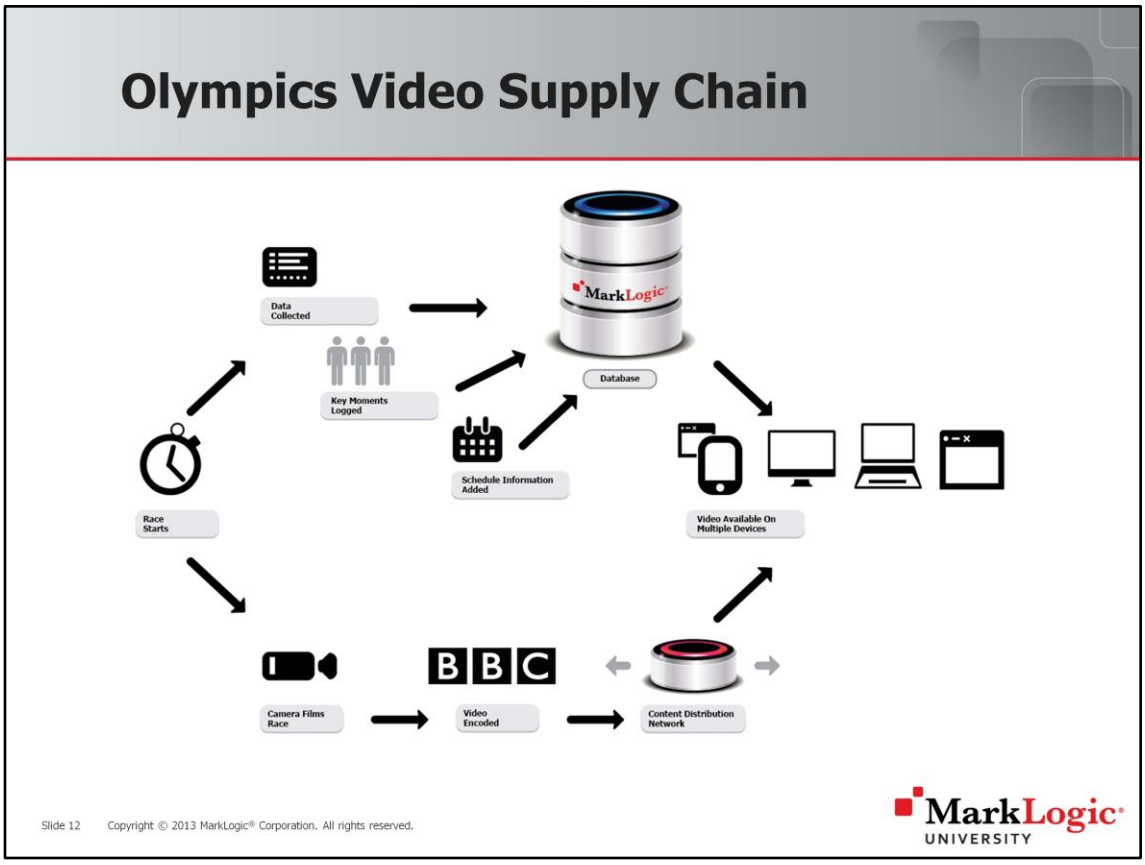
- The last common use case we'll discuss is related to Content Delivery
- Often what we see here is some app which connects to a large number of data sources
    - This is a lot of custom work
- Then along come another product and another. We end up with a rats nest of connections between applications and data sources
- The problem gets worse every time a new data source or application is added.
- By this point I think you can guess what the solution looks like



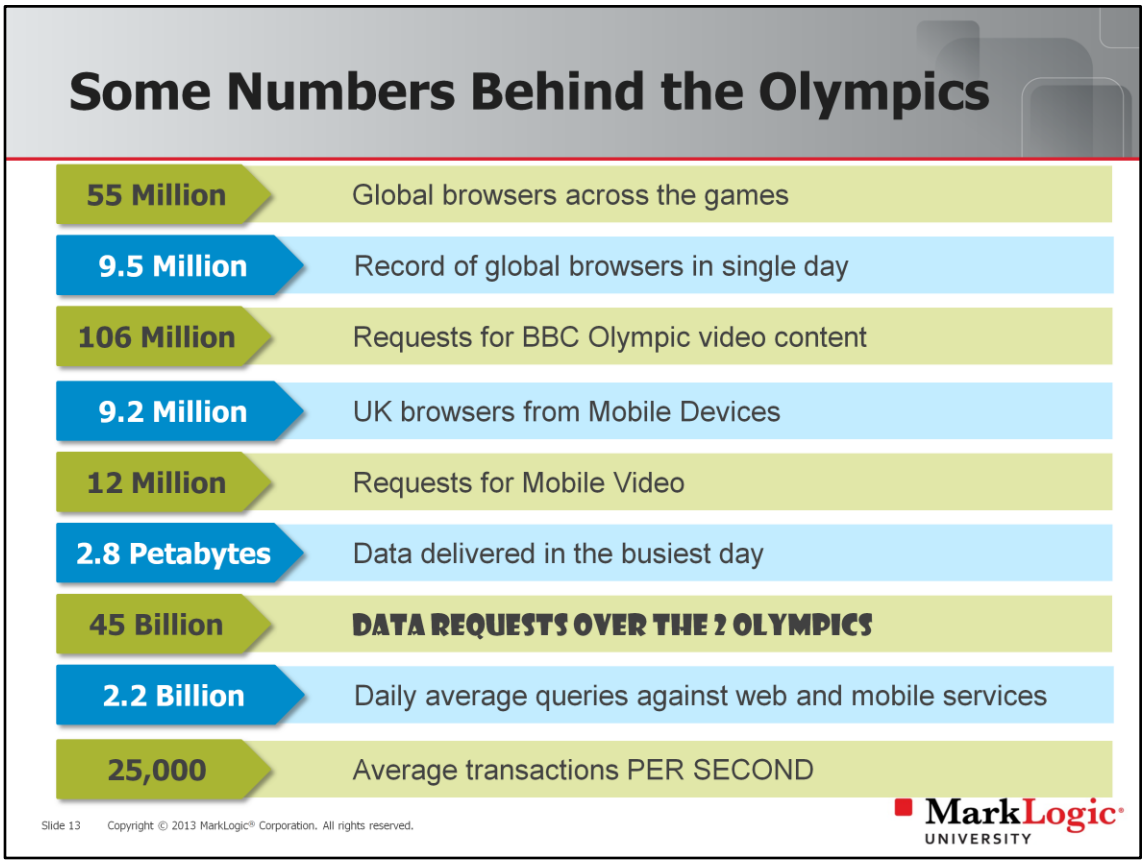
- We load data from disparate sources as-is
- AND we can publish to many different device types and form factors in real time



- With our Enterprise NoSQL solution you don't have to compromise
- You get all of the Trusted features that enterprises need, …
- Much greater agility than other NoSQL technologies provide, …
- And the power that's available in both SQL and NoSQL



BBC delivered over 10,000 pages of dynamic web content with just their normal staff of 14 journalists.



So how big was big data at the Olympics. Here are some stats for the BBC…

REST & JAVA APIs
MarkLogic is REST Server

REST interface to MarkLogic Server

Foundation for language-specific API's

- Full-text and faceted search
- Access to value indexes and aggregates
- Document management for XML, text, binary, and JSON
- Separate read, write, and administrative access
- Administrative services for managing API configuration
- Extensibility with custom URL endpoints
- Scriptable and UI bootstrapping

Pure Java API to MarkLogic

Benefits

- ✓ Leverage existing skills
- ✓ Speed application development
- ✓ Address your application backlog
- ✓ Reduce development costs
- ✓ Improved reuse and efficiency

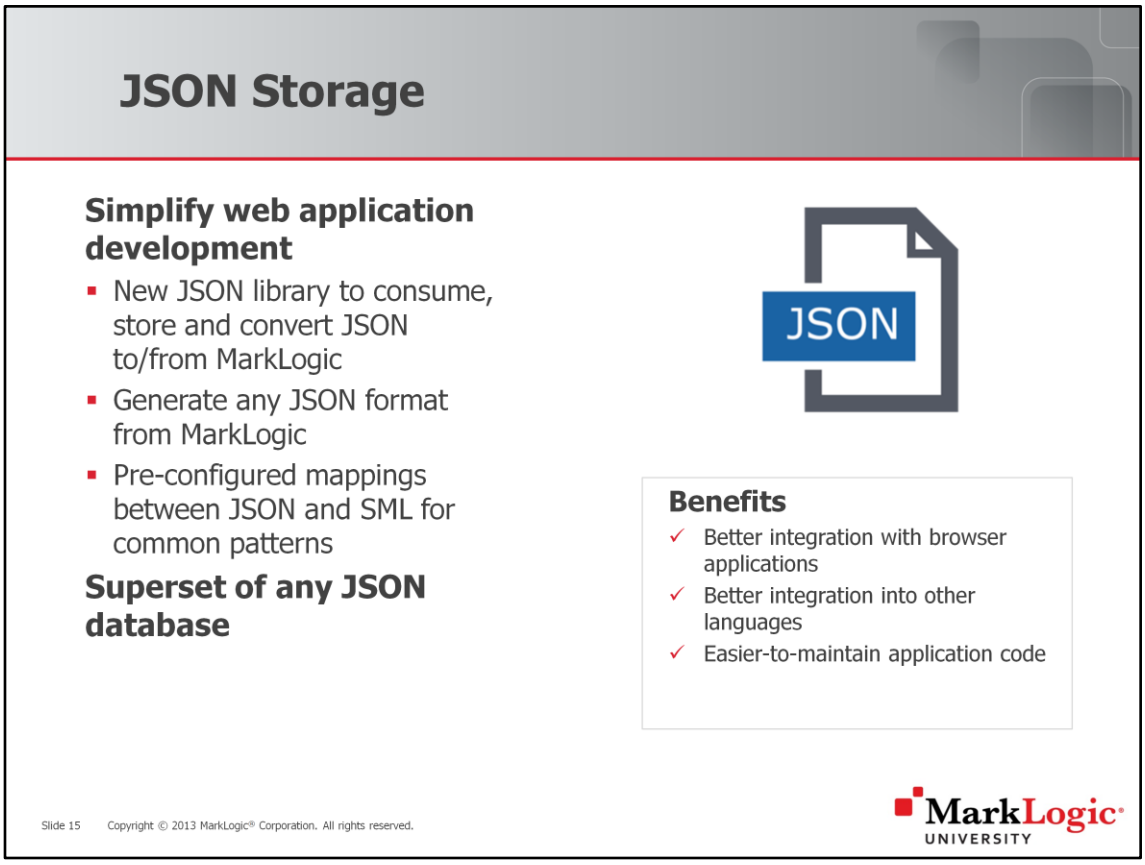
MarkLogic UNIVERSITY

Leverage MarkLogic's optimization for Big Data on traditional programming tools
- Schemaless
- Indexes
- Structure-awareness

Focus on advanced features and innovation, not infrastructure
Leverage Java ecosystem (IDEs, debuggers, etc.)
Java developers able to quickly learn MarkLogic with the simplified API
Reduce infrastructure/maintenance costs by MarkLogic capabilities into existing Java-based application infrastructure

(What this means: The MarkLogic Java API is a pure Java API to MarkLogic. It allows programmers to take full advantage of the Java platform.)

Benefit from MarkLogic's optimized Big Data capabilities without adding infrastructure or maintenance costs



With the new JSON support, developers will be able to get JSON into and out of MarkLogic in an XML-centric application using XQuery and XSLT. Those working primarily with JSON in non-MarkLogic middleware (J2EE, .NET, PHP, etc.) will be able to leverage MarkLogic for search and CRUD using a JSON data model.

It's now easier to build web applications

Visualization Widgets

Quick Development of Visual Interfaces

- Easily build or enhance web applications using MarkLogic widgets
- Simplify front-end architectures with pre-configured visualizations and data access helpers

Explore New Ways to Look at Data

- Gain more insights from data sets with visual representations
- Charts, Maps, Search, Results, Sidebar with facets

Benefits

- ✓ Simplifies development
- ✓ Speeds time-to-deployment
- ✓ Easy access to powerful geospatial and aggregation capabilities
- ✓ Scalable, event-driven, encapsulated

Slide 16

MarkLogic UNIVERSITY

Customers love visualization (and equate “analytics” with visualization), and it’s important to be able to develop applications that use visualizations – these widgets make it much quicker/easier for developers to build cool applications that are really useful for their end users.

Data visualization is requirement for most modern analytic applications. Visualization allows users to discover the shape and dimensions of data and quickly assess trends and patterns over potentially large data sets. It also allows users to explore data more naturally than something like a query language for many use cases and non-expert users. Consumer applications are increasingly including sophisticated visualization into the common vocabulary of the web. This has pushed enterprise applications to become more responsive, interactive, and visual as users have come to expect.

BI Tools Support

Big Data Analysis for mere mortals

- Perform sophisticated analytics on ALL your data, in real-time

Leverage existing tools avoids added expenditures

- No additional training
- Reuse existing report templates
- No custom integration or code required

Simplify IT infrastructure

- Integrated through ODBC – tested on IBM Cognos & Tableau
- No need to spend resources on extracting data to a data warehouse
- Real-time access to your operational database – gain competitive advantage

Benefits

- ✓ Makes your data more accessible across your business
- ✓ Faster results
- ✓ Faster ROI

Slide 17

Tested for Out-of-the-box integration with the top BI tools: IBM Cognos and Tableau.
Possible future support for SAP BusinessObjects

Content Pump

Simplifies moving data into, out of, and between MarkLogic databases

- Greater reliability from commercialized code
- Leverage existing infrastructure to import new data

Enables loading of large data sets in parallel

- Available to leverage Hadoop to parallelize loading

Bypass bottlenecks, load directly to data nodes

Benefits

- ✓ Less custom code = lower costs
- ✓ Lowers cost and complexity
- ✓ Speed time to deployment
- ✓ Faster performance

Slide 18

MarkLogic UNIVERSITY

Scriptable, command-line tool that simplifies moving data into, out of, and between MarkLogic databases

- Import data from the file system, Hadoop (HDFS), or standard input
- Export data from a database to the file system or Hadoop
- Archive data from a database, including its permissions and metadata, to the file system or Hadoop
- Copy data between two MarkLogic databases in different clusters

Users need to be able to efficiently get data into and out of MarkLogic databases – including to and from HDFS (Hadoop Distributed File System).

We currently have a number of mechanisms to do this (Information Studio, XQuery APIs, RecordLoader, etc.), but Content Pump fills in the gaps – and is officially supported by MarkLogic

Hortonworks Hadoop Distribution

Certified bundle of the Hortonworks Data Platform and MarkLogic Connector for Hadoop

- Essential components to run Hadoop in production
- Built-in installation and provisioning tools
- First-line technical support from MarkLogic (second-line from Hortonworks)

Run Hadoop MapReduce on MarkLogic data

Benefits

- ✓ Easy interoperability with Hadoop tools
- ✓ Reduced development costs
- ✓ Eliminates integration costs
- ✓ More speed
- ✓ World-class support

Slide 19

MarkLogic UNIVERSITY

- Distribution of Hadoop combines Hortonworks Data Platform with MarkLogic and enterprise-grade support
  - Easy: interoperability with Hadoop tools
  - Fast: leverages MarkLogic's indexes and distributed architecture for performance
- Unique benefits of combining Hadoop with MarkLogic
  - Wider range of analysis on Big Data
  - Real-time interactivity provides immediate answers plus drill-downs into specifics
  - Offline, batch analytics for answers where real-time is not necessary
- Reduced development costs of building out a Big Data platform
  - Out-of-the-box capabilities: search, alerting, geospatial, and transactions
  - Eliminates integration costs of stitching together tools
  - World-class professional support

Unit 1: Applying the Learning Objectives

- List the three core components of MarkLogic Server
- Explore example use cases and applications
  - Exercise 1
- Describe Enterprise NoSQL