# DATASTAX

> "DataStax Enterprise and Microsoft Azure help us collect, validate, and distribute almost a full terabyte of content and convert it to value for our customers."
>
> *Graham Lammers, Director*
> *Product Development and Delivery*
> *IHS Markit*

**MASTER DATA MANAGEMENT**

# 500 MILLION COMPONENTS: IHS MARKIT CONVERTS DATA TO VALUE WITH DATASTAX

When IHS Markit found itself managing a digital catalog of more than 500 million components with a disparate set of database technologies, it worked with DataStax to unify and automate its electronic parts database into a single system, the Parts Content Factory. IHS Markit used DataStax Enterprise, the leading multi-model NoSQL database along with the Microsoft Azure cloud platform—to give its database the flexibility, high performance, and scalability that the company needed. With new search, analytics, and data-visualization capabilities, IHS Markit can keep the Parts Content Factory humming, while its customers can find the parts—and the insights—they need.

---

A digital catalog with more than 500 million electronic parts and components—for everything from smartphones to satellites—takes a tremendous amount of work and resources to build, manage, maintain, update, and productively use. That might explain why IHS Markit calls its database the Parts Content Factory.

From daily operations to global strategy, IHS Markit provides insight and services to Fortune 500 enterprises across business lines that include aerospace, automotive, energy, and technology. The IHS Markit technology line delivers full lifecycle services, analytics and data—from silicon fabrication and consumer-connected devices to enterprise IT, medical devices, manufacturing automation, and a teardown and cost modeling practice.

The Parts Content Factory provides mission-critical support across the entire IHS Markit technology line and for a customer base that includes leading electronics companies and a diverse set of roles and practices such as component engineering, quality engineer, design engineer, supply chain management, marketing, global data managers, regulatory compliance officers, and procurement. They use BOM Intelligence tools supported by the Parts Content Factory as a reliable, objective source to monitor component status, part change and end of life notifications, and regulatory compliance for RoHS, REACH, conflict minerals law, and U.S. Export regulations.

When necessary, IHS Markit customers can use BOM Intelligence to quickly and accurately identify alternate component sources, optimize spend on electronic components, maximize part reuse, identify preferred manufacturers, select component parts with a life cycle that is appropriate for their product, leverage alternate parts for optimal use, pricing, and status, and track and maintain regulatory compliance across the globe.

So the IHS Markit Parts Content Factory has to be powerful, fast, and dependable. That's why IHS Markit uses DataStax Enterprise to give its Parts Content Factory the flexibility, high performance, and scalability it requires to meet the complex needs of IHS Markit customers.

## Challenges

- Maintain fast performance for a digital catalog of 500M IT components
- Difficulty searching quickly through complex attributes and relationships
- Manage disparate and brittle legacy database system

## Objectives

- Unify and automate electronic parts database into a single system
- Improve and accelerate search experience
- Decrease time required for IHS Markit knowledge workers to update the system with new information
- Create better customer experiences and meet complex customer needs

## Solution

- Selected DataStax Enterprise for flexibility, performance, and scalability
- Worked with DataStax to build data model to help unify and streamline database into Parts Content Factory
- Uses search and data-visualization capabilities into DataStax Enterprise to generate fast query results and real-time business analyses
- Runs DataStax Enterprise on Microsoft Azure cloud platform

## Results

- Reliable performance increases revenue opportunities
- Database team can do more in less time at lower cost
- Focus on business innovation instead of system maintenance
- Leapfrog capabilities, deliver more value, keep pace with business

## A factory in the cloud

IHS Markit grew its parts business in several acquisitions, and processed all the associated data with a disparate and disconnected mix of relational databases. "We wanted a faster, more unified, more robust system," says Graham Lammers, Director of Product Development and Delivery at IHS Markit. "We knew that to turn a list of 500 million parts and components and several billion pieces of related content into actionable information and insight, we would need a highly scalable framework."

IHS Markit wanted to unify the database in a single system, the Parts Content Factory, but the company saw that the long-term demand and growth requirements of its parts business were greater than it could manage cost-effectively with a relational database. It examined NoSQL distributed databases and determined that the Apache Cassandra™ masterless architecture would work best with its datasets. With its massively linear scalability, ring-based clusters, wide columnar, and CQL interface, DataStax Enterprise, built on Apache CassandraTM won out over the competing NoSQL technologies such as Apache™ Hadoop® with Apache HBase™. With DataStax Enterprise (DSE), IHS Markit can leverage critical database features and expert support, all day, every day.

"We reached out to DataStax for guidance to help us choose the data model that fit our needs," says Lammers. "We chose DataStax Enterprise for its search, analytics, and data-visualization capabilities."

With DataStax Enterprise Search's deep integration of Apache Solr™, IHS Markit can easily optimize DataStax Enterprise specialized queries across a diverse set of complex attributes to generate fast and reliable results. With simple linear scalability in DataStax Enterprise, IHS Markit can accelerate query execution by just adding resources to any completed Solr schema.

Additionally, with DataStax Enterprise Analytics, powered by Apache Spark™ big data analytics processing engine, IHS Markit can build dashboards for data-table visualizations and real-time business analyses. DSE Analytics will help support data science and machine learning that IHS Markit will use to automate content classification and entry, an advancement that will literally cut days out of the intake process, getting data updates to market in less time—in many cases weeks ahead of what can be accomplished today.

IHS Markit has already used DataStax Enterprise Analytics to help eliminate many ETL responsibilities from existing databases. By reducing, simplifying, and eliminating most of the legacy ETL processes through DataStax Enterprise, IHS Markit can not only drive faster updates to market but, more importantly, can radically reduce the overall maintenance costs and complexity of back-end systems.

IHS Markit runs DataStax Enterprise on the Microsoft Azure cloud platform to scale up and down as necessary and because Azure resources performed as well as dedicated physical machines in early proof-of-concept studies for the project.

## Data converted into value

Almost 350 IHS Markit knowledge workers use a dashboard to collect parts updates and inserts through web scraping and manufacturer data feeds. Depending on the category of parts, updates and new parts can scale to millions of parts per week. IHS Markit knowledge workers also search through hundreds of millions of parts within the Cassandra database to find and compare part attributes and relationships for potentially every entry. New parts can also be manually entered with all relevant data using an ASP.NET WebAPI / Angular JS application.

Customers sign in to the IHS Markit BOM Intelligence service to manage electronic components from design through production. It is more critical than ever for IHS customers to stay up to date on obsolescence, counterfeit, and non-compliance risks in their supply chain. An effective BOM analysis system enables them to increase the pace of new product introductions, avoid costly production interruptions or product redesigns, and improve sustainability over extended service lives.

## Agility to grow, consistent performance, and more sales

With DataStax Enterprise, IHS Markit can operationalize the growth of the Parts Content Factory to support temporary spikes in usage or business growth, whether it's organic or acquisition-based. "With DataStax Enterprise, we can keep pace with the business without having to re-architect the system every three years," says Lammers.

With a powerful, stable, high-performance search and analytics capability, the Parts Content Factory performs reliably from one query to the next and from one user to the next. Customers find the parts—and insights—they need, while IHS Markit increases revenue opportunities and employees do more in less time and at lower cost.

## About DataStax

DataStax, the leading provider of database software for cloud applications, accelerates the ability of enterprises, government agencies, and systems integrators to power the exploding number of cloud applications that require data distribution across datacenters and clouds, by using our secure, operationally simple platform built on Apache Cassandra™. With more than 500 customers in over 50 countries, DataStax is the database technology of choice for the world's most innovative companies, such as Netflix, Safeway, ING, Adobe, Intuit, Target and eBay. Based in Santa Clara, Calif., DataStax is backed by industry-leading investors including Comcast Ventures, Crosslink Capital, Lightspeed Venture Partners, Kleiner Perkins Caufield & Byers, Meritech Capital, Premji Invest and Scale Venture Partners. For more information, visit DataStax.com or follow us @DataStax.