> "Talend and Snowflake are core components of the platform we've built to deliver top-tier healthcare data and analytics services."

Sven Junkergard, CTO, DRG

DRG
DECISION RESOURCES GROUP

INDUSTRY

• Healthcare / Life-sciences

INFORMATION

• HQ: USA

• 1,000–1,500 employees

USE CASE

• Healthcare Business Intelligence

CHALLENGE

• Building a specific kind of big data platform

TALEND PRODUCT USED

• Talend Cloud Real-Time Big Data

RESULTS

• **150% more productive** with same costs

• **Onboarded 100-plus TB of data in** 3 months

• Supporting **5x more users** at same cost

Finding actionable insights in healthcare big data

Founded in 1990, DRG (Decision Resources Group) has become the premier provider of healthcare analytics, data and insight products and services to the world's leading pharma, biotech and medical technology companies. DRG's in-house team of expert healthcare analysts, data scientists, and consultants provide information, analysis, and insight on critical issues to the healthcare industry worldwide. It offers solutions across various disease areas; syndicated and custom market research; research publications; and advisory and consulting services on trends, emerging developments, and market potential.

Building a specific kind of big data platform

"Our clients are facing ever-increasing competitive pressures driven by the pace of healthcare industry innovation and discovery," says Sven Junkergard, CTO, DRG. "Our clients also have to deal with constantly evolving healthcare markets and regulations, decreased healthcare budgets, increased consumer power and choice, and a high degree of M&A. Undeniably, we are in the midst of the most disorganized and misunderstood healthcare landscape ever."

The current landscape also includes an explosion of healthcare consumer and system data. There's an arms race emerging in the world of market research, as big data promises to revolutionize the business of pharmaceuticals, from R&D to commercialization and beyond. Because more data than ever before is freely available, primary research can provide only part of the picture of global healthcare and disease markets. Therefore, to remain a market leader and respond to increasingly specific requests from its clients, DRG needed to organize and combine a number of data sources, analyze information faster, and deliver more insightful analysis via several channels.

"We primarily help our clients do three things," says Junkergard. "They are: understand their markets, optimize patient access to treatment, and engage their customers." We recognized that to do that better than our competitors, we needed to be able to manage and derive insights from petabytes of data. That data includes the research data we own and manufacture, 'real-world data' from millions of electronic health records [EHRs], healthcare claims and medical records, and data from clients. We knew our existing IT infrastructure could not scale up to handle that volume of big data, so we set out to create what we call a Real World Data Platform."

Junkergard says that a key consideration in evaluating possible solutions for the platform included finding an SQL engine that works with big data, because the company's analysts were already familiar with SQL. The new engine also needed to meet specific performance requirements. "A big part of our infrastructure is already in the AWS cloud," he says, "so we felt that would be the right place to deploy the rest of our data platform. In a nutshell, we knew we needed a data engine that was an ACID-compliant distributed SQL platform, could support multi-terabyte data volumes, and was hosted in the cloud. What we didn't know was if such a platform existed or not."

Why Talend?

After evaluating alternatives, DRG selected Talend and the Snowflake cloud data warehouse as the foundation of its new Real World Data Platform, and worked with Talend Professional Services to implement the cloud solution. Other key components of DRG's new platform include AWS S3, Spark for advanced machine learning, Apache Parquet format, and Tableau for analysis. In essence, the company built the very platform it was looking for.

"A key reason we chose Talend is that it's compatible with big data," says Junkergard.

Junkergard says another deciding factor is that Talend's infrastructure is in the cloud, which gives DRG scalable compute performance without complications. "Because of that," he says, "our developers build data integration workflows, without much coding involved. In fact, our data platform has helped us scale in a short period of time from serving a handful of users to serving more than 200."

DRG also relies on Talend for data governance. "Healthcare data is quite messy," says Junkergard, "and ensuring data quality is a job that never ends. For instance, standardizing values from different labs in different systems is a challenge, and we do it with Talend." He adds that Talend helps ensure data quality via extensive upfront data cleansing, reference data mapping, and a data quality framework that tracks errors and shifts in underlying data. Junkergard says Talend has helped DRG master its data and build a trusted single source of the truth for departments companywide.
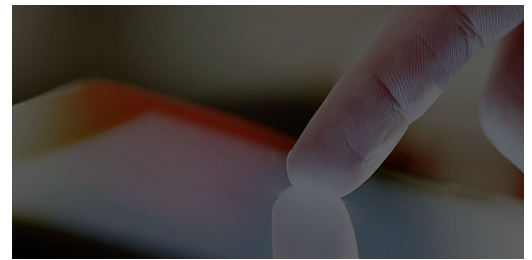
"With all its components and connectors, Talend plays a pivotal role in our architecture by connecting all the pieces with minimal coding," says Junkergard. "Talend is keeping us close to the cutting edge in healthcare data management by integrating new technologies as they become popular, without the need for us to re-tool."

Tracking the patient journey, the influencers and segmenting the market

DRG makes use of its data platform—whose data repository now contains several hundred terabytes of data—to support a range of use cases. One of the fastest growing segments of DRG's business is patient and market segmentation. DRG is building on syndicated offerings, which analyze the current state of care at the macro level, to offer more specific research customized to precise categories of patients and targeted therapies. "There's increasing innovation around treatments and personalized medicine," says Maria Cipicchio, Director of Campaign and Segment Strategy, DRG, "so the pools of prospective patients are getting smaller, which requires ever-more-precise targeting. With much bigger data sets and faster, more in-depth

analytics, we can deliver analyses to our clients that help them allocate their resources geographically and by their most important patient segments."

Another use case is tracking the patient/customer journey and buying process, which includes analyzing steps in a cascade of treatments, the role of the patient experience, and more, so that clients can better map their assets to steps in the journey.

A third key use case is commercial targeting for physicians and health systems. "We can map the influence of an individual physician and offer insights to our clients about how to focus their sales efforts," says Junkergard. "Healthcare data isn't always stored in a structured way, so we have to analyze millions of records and group data to create a picture of a sphere of influence. Talend drives all that—our algorithms get coded and executed in the Talend environment."

According to Junkergard, perhaps the biggest benefit of the Talend/Snowflake infrastructure is that it has enabled DRG to cost-effectively centralize, navigate and rapidly analyze high-quality data. That capability, in turn, has made it possible for the company to enter, scale up, and succeed in the fast-growing analytics area of healthcare market intelligence. "Our data platform has enabled us to support a very high growth rate with a smaller overhead investment than other solutions would have required," says Junkergard. He says DRG's IT staff can now serve a large number of users who get their answers faster, which enables them to deliver products and services to market sooner.

"Talend and Snowflake provide the necessary infrastructure to accommodate and sustain massive growth in our data assets and user groups over time, enabling us to maximize the value of our investments," concludes Cipicchio.

> "Our team searched for the best possible technology to support our new reality, and we decided to move forward with the combination of Snowflake and Talend. It has been a great success story for us so far. Today, we're supporting many times more data and users than before our move."