Datalytyx
Making Information Work

wejo transforms real-time big data into actionable insight in a connected world

A real-time Modern Data Architecture founded on Talend, Cloudera and AWS technology and delivered using Datalytyx expertise is enabling a fast-growing automotive industry start-up to exploit real-time big data streams and create innovative new journey-based products

wejo

Client

Pioneering provider of journey and location-based 'smart driving' services

Challenge

Create a powerful and scalable environment to enable wejo to "revoluntionise the automotive industry" through products and services based on big data, telematics and mobile technology.

Solution

'Quick Start' delivery of an AWS-hosted MDA with a Kappa design pattern: Talend data integration and management, Cloudera for Hadoop distribution, Apache Spark streaming, Apache Kafka/Talend for data pipelines.

Benefits

An extensible, scalable, secure and flexible platform; generating commercial value from real-time Big Data streams to consistently deliver more for clients and drivers while building revenues for wejo.

Challenge: create a powerful and scalable environment to drive business success

A pioneering provider of journey and location-based 'smart driving' services, wejo aims to revolutionise the automotive industry through leading-edge business-to-business-to-consumer services created using a combination of big data, telematics and mobile technology. "We work in the connected car space," says Steve Pimblett, COO. "Motor manufacturers around the world are asking 'what can we do with all the data we're collecting? How can we use it to help drivers; how can we commercialise that data?'" With an approach based on a secure and highly-scalable Level 5 data exchange platform, wejo's own-brand, white label and bespoke app developments will use rewards-based models and captured journey data to deliver actionable business insight for OEM and insurance clients, helping them connect with customers in smarter ways.

Solution: a modern data architecture (MDA) founded on Talend, Cloudera and AWS

"We asked our Talend account manager to recommend the best consultancy," Pimblett continues. "We met with Datalytyx and the chemistry, the relationship, was there from the start. The Datalytyx people were commercially *and* technically savvy, and also demonstrated emotional intelligence. Their Quick Start approach fitted our needs perfectly. The whole Datalytyx approach was pragmatic in terms of wejo being a start-up business with budgetary constraints: a willingness to change scope as needed, and being as flexible on budget as possible." Datalytyx delivered an extensible, scalable and secure MDA hosted using Amazon Web Services, Cloudera for Hadoop distribution, Apache Kafka and Talend for data pipelines. The Kappa design pattern means all data is ingested in streams rather than batches, and can be delivered to clients as streams.

"Talend was the starting point," Pimblett says. "I'd built a similar architecture at my previous company but we tried to handcraft it. It was good but needed a lot of maintenance and overheads. At wejo, we wanted a full management framework: operational management and dashboard, security out-of-the-box, rapid deployment out-of-the-box. From a data integration platform perspective, Talend is the market leader, and it also matched our in-house skills set: no need to handcraft integration, or to send integration jobs overseas then perform additional work in the UK. Talend

> "We needed an architecture able to handle massive data points on a substantial scale, generated on a continuous basis, with GPS co-ordinates data pushed every second or millisecond. Our real-time architecture had to capture that, and enable wejo to turn all that data into insight and actions."
>
> **Steve Pimblett**
> **COO**

also means lower cost of ownership. It's very easy to train people, and scale up your internal team." Source systems for the MDA include existing environments in wejo's data estate and acquired third-party data feeds.

"We knew it would be a cloud solution, not least so we could scale at will," Pimblett continues. "We had some operational services running in Amazon already, so it was logical to follow that route for the MDA - a natural fit. AWS offers the security and resilience we want, along with managed services as you need them – meaning it's also a good fit for a start-up, giving access to lots of added services." Cloudera was recommended by Datalytyx: "We see our Trust Framework as a unique selling point; it's essential so we can embed the four pillars of Security, Privacy, Integrity and Anonymity in our architecture, for governance and compliance. Based on advice from Datalytyx and our own research, we were confident Cloudera would be the best Hadoop distribution and support our streaming requirements with the use of Apache Spark Streaming. Our business is founded on being a data platform and data aggregator, which means handling different data types, ensuring data quality, and effectively managing and leveraging all of that data. So our MDA was about security, data quality, and data management."

A real-time approach: generating value from Big Data

"Because of the nature of our industry, with data streaming from vehicles, wejo had to be a 'real-time' rather than 'overnight' business in terms of segmentation and turning data into insight into action. For instance, sending a local retail offer to a driver based on their location: as they apply their handbrake in a specific car park. We need to push those targeted notifications and alerts in near real-time. And if there's an accident, we want to alert local services and an insurer's contact centre to check that driver and passengers are okay. It has to be a streaming approach in real time. So the Kappa architecture was a perfect fit." Kafka message queuing is another key element, passing data between internal elements of the solution. The result is a cutting edge architecture that will enable wejo to reflect changes in driver behaviour in near real-time and to offer new, highly differentiated services. "Open source, well established and highly effective for queuing, Datalytyx recommended Kafka, and it was another perfect fit," Pimblett says.

Architecture Design work by Datalytyx included workshops to understand requirements in-depth and to draw the target architecture, leading directly into deployment. The Datalytyx Quick Start engagement focused on configuring infrastructure and technology platforms: setting up AWS and installing Talend and Cloudera, through a development, test and live environment, all with appropriate security and ready to run. Datalytyx is also supporting the build of new data products. The first is a Journey Streaming 'black box' system for trucks Europe-wide: a real-time data stream enabling traffic alerts for client companies. Pimblett adds, "We have the

Data infrastructure so we're now focusing on *Insight* and *Action*." Over time, the MDA will enable other data feeds including weather, social media and event data to integrate with core driver and vehicle data, adding even more value to wejo's propositions: the opportunity is to consistently deliver more for clients and drivers - and build revenues for wejo.

About Datalytyx

Datalytyx is a leading UK provider of big data, digital transformation, governance and analytics solutions, all delivered as a service from the cloud. Our client-focused approach delivers rapid gains and sustained ROI, meets requirements to store and manage huge data volumes, deploys powerful business-focused analytics, and digitises essential business content and processes. The trusted management information that results enables more effective and more profitable decision-making. Our diverse client base includes AstraZeneca, Nissan, ABN AMRO, Surrey County Council, EasyJet, Allianz, Unisys, Capita, Eversheds, Veolia Environment and UK local government. We also partner with other leading providers to deliver best-of-breed solutions, including Microsoft, K2, Talend, Tableau, HP Vertica, KnowledgeLake and Hortonworks.