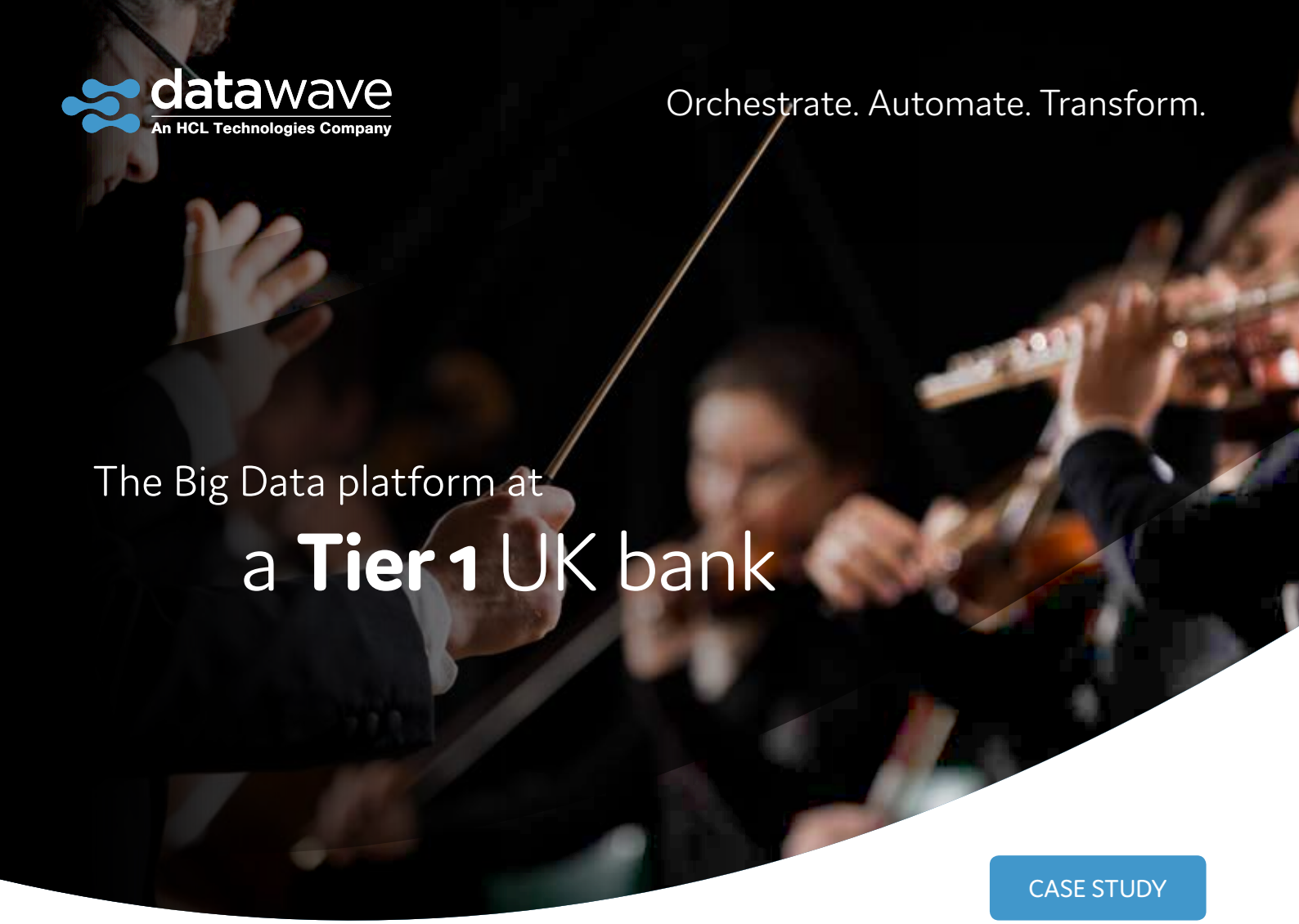
datawave
An HCL Technologies Company

Orchestrate. Automate. Transform.



The Big Data platform at

a **Tier 1** UK bank

CASE STUDY

A Tier 1 UK bank has developed a Big Data strategy which it calls "personology" in an attempt to reconnect with customers. The bank is using data analytics to create a personalised approach to customer service. The philosophy is one of the developments of the 800-person strong analytics department, created as part of a £100 million investment in analytic skills and technology across the organisation.

Key facts

- Personalised approach to customer service using data analytics
- 800 person strong analytics department
- £100 million investment
- Data assets acquired across 160 major systems
- Handles over 1 Terabyte of data on a daily basis

"In comparison to the last release under the old model, we have delivered twice as much data, in half the time, at 25% of the cost."

Quote from Tier 1 Bank

If the Bank was to succeed with a strategy that moved the focus from data-driven marketing to restoring the trust and feeling of support that bank customers expected in the past, then they needed to understand their customers better. The Bank identified that it could achieve this by making better use of the huge richness and depth of customer data that it holds to personalise how they deal with their customers and their needs.

As well as partnering with Datawave to deliver their Big Data success, the bank has invested in open source big data technologies such as Cloudera Hadoop and Cassandra and Big Data-driven service providers – such as CRM specialists Pegasystems PEGA +0% – to drive their analytics operations.

The use of the Datawave platform at the bank has been a key element to implement this strategy; Datawave has enabled the Bank to rapidly acquire significant data assets in a rolling programme across over 160 major systems and many more medium level applications within the Bank; each is made up of tens, hundreds, and in some cases, thousands, of separate files and tables hosted on a variety of technology platforms including Oracle and the Bank's backbone mainframe environment.

All of these systems are rich sources of customer data, but often relevant data is spread across transactional systems that were developed years or even decades ago; some were acquired through acquisition of other Banking groups, others were built from the ground up and have been much modified over the years. For the Bank's Analytics teams rapid and properly governed access to this data was crucial if the Bank's objectives were to be met. The bank identified Datawave's automated delivery platform as the solution to meet the speed, scale and complexity of this data challenge and to provide a robust data provisioning framework that could meet the operational and data governance standards that are essential to build confidence and security in the Bank's Big Data platform.

What has the Bank achieved?

The Datawave platform has enabled the Bank to deliver an enterprise scale platform that provides:

* Automated data provisioning using Spark and orchestrated by our UI-driven framework incrementally loads over 1 Terabyte of data on a daily basis to the Bank's ever-expanding HADOOP platform based on Parquet
* HIVE tables and views are automatically refreshed each day
* Provision of Entity and field level documentation that supports an always-growing data catalogue enriched with operational metadata that records where data originated from, when it was last refreshed and if data has been dropped
* Implementation of critical data security policies through powerful data masking rules that are easy to apply.

How the Datawave platform delivers the Bank's needs

1. Automated Development

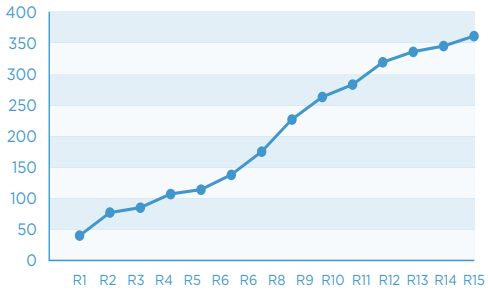
Firstly, automation is seen as essential if the Bank is to achieve success at scale since despite the advances in technology and the development of powerful Big Data programming languages, provision of data at scale had relied upon teams of developers with specialised skills coding data-provisioning routines. The Bank identified that a step-change in productivity required the automated development of these data-provisioning processes. Datawave's platform builds and executes the provisioning process in a single step, taking the data from a staging area straight to the HADOOP platform in a single step that also delivers HIVE tables and Views. Just as importantly, Datawave enables this without the user having to write a single line of code which overcomes the skills shortage challenge to delivering Big Data at scale.

2. Data Governance

Secondly, Data Governance was identified as essential if the Bank was to deliver its core objective to understand and respond to customer's needs. Having the data was just the first step; without knowing what data the platform contains, what it means, who uses it, what it is for, and how good it is then the data would not help to deliver the fundamental needs of the Analytics community. Data security matters as well; the option to mask sensitive data should be available across the platform.

3. Operational Robustness

Thirdly, operational robustness was essential: the Bank needed the ability to orchestrate data provisioning processes by defining how processes are combined and scheduled to deliver the right data to the right place at the right time underpinned by robust error-handling, job recovery and a rich set of operational metadata visible to technical and business users.



The Datawave platform has enabled the Bank's Big Data platform to become an operational necessity that is in constant use. From a start in mid-2016, the Bank achieved the 15th release milestone for the platform at the start of 2018. This brought the total number of separate source systems on the HADOOP platform to over 400, each of which is made up of tens, or hundreds of source tables or files. The table to the left shows this rapid rate of growth:

datawave
An HCL Technologies Company