

profusion

Democratise data across your company

Eucalyptus (*modern data infrastructure*): provide a powerful and unified analytics environment for big data

Data
influenced
journeys



Democratise data across your company

Challenge

Data from Adobe Campaign, Kickdynamic, Dynmark and other data sources was difficult to access.

In September 2013, Profusion's data science team responded by designing a new analytical database (*ADB*) that merged data from multiple siloed data sources into one unified data warehouse.

The database empowered data scientists to query all multiple data sources using SQL. Data analysts were able to use the new infrastructure to produce detailed marketing reports in one hour instead of eight. This marked a radical shift for the business, enabling quick insights for our teams of marketing consultants and campaign managers.

Eucalyptus backend

We are moving towards a world where we must deal with increasingly large datasets for the work that we do. We use Hadoop, Spark, Zeppelin and Hue to provide a powerful and unified analytics environment for big data.

Hadoop

Hadoop is a data infrastructure designed for parallelised computation. The advantage of using a Hadoop cluster over a server with a traditional SQL based data warehouse is this: we can organically grow and scale it by adding one Hadoop node at a time.

Spark

Using Hadoop, we are empowered to use parallelised computation. However, writing parallelised algorithms is very time-consuming, complex and challenging. Using Spark, we can hide away a lot of that complexity, make it very easy for data scientists to harness machine learning libraries optimised for parallelised computation. Spark provides a unified framework where querying and analysis can happen at the same time.

Notebook environment: Zeppelin and Hue

Using Zeppelin and Hue creates a user-friendly notebook environment where Spark and Hive codes can be written and shared with colleagues. Here it is much quicker and easier to explore data and develop code than it is at the command line.

Eucalyptus semantic layer

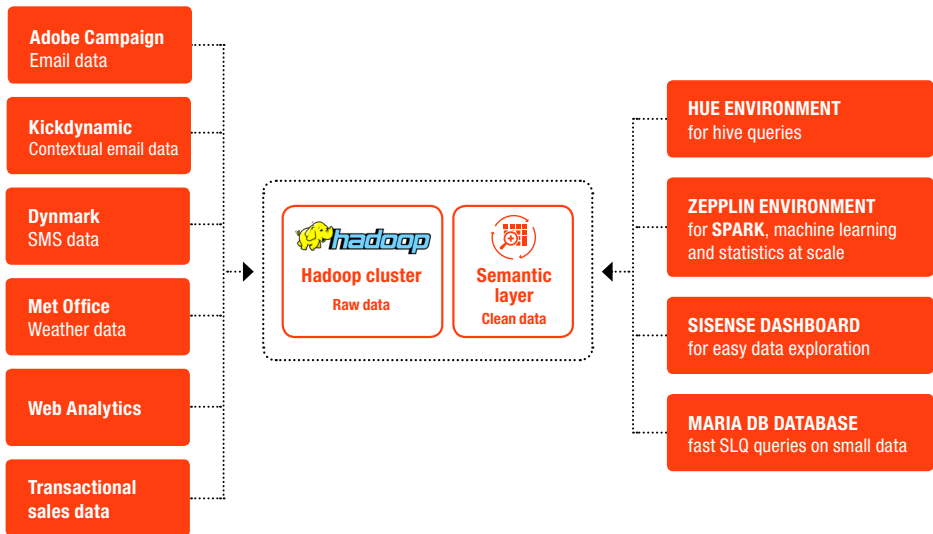
Currently it takes three to four months to train new data scientists to use and query the analytical database. Within Adobe Campaign, the data table structure is very complex. Adobe has been designed for users to log and store data efficiently, but not to understand and query it easily.

Our data science team has designed new clean and clear data tables known as the semantic layer. This allows newly recruited data analysts and data scientists to get to grips with the data at speed and produce valuable insights soon after they join. They need relatively little training to do so.

Eucalyptus dashboard

Using the clean data tables, this data is fed into a dashboard to empower marketers, campaign managers, email developers and third party stakeholders. Anyone can explore and ask questions of our marketing data, without a data analyst or data scientist. Key stakeholders can explore and interpret the data in real time to inform commercially-led decisions around time-critical client or customer needs.

Eucalyptus data infrastructure



Wired to solve problems

We're a clever team of specialists in data engineering, BI, data science and multi-channel marketing.

Blending real 1:1 communications, AI, machine learning and human intelligence, we connect people with your brand.

We were the original UK partner for Adobe Campaign. We've nearly 20 years' experience in data management, automation and customisation of enterprise data management and marketing other platforms.

Fuelled by inquisitiveness, forward thinking and excitement at the advancements of technology, we're entrepreneurial, adventurous and we make stuff happen.

We help clients get under the skin of a human being's emotional narrative – their story – and see what's important to each customer.

Our clients trust us for our responsive, personal and upfront approach to help them create highly targeted, intelligent communications.

Stay connected to discover more inspiring guides and stories within our 'knowledge think tank'.



**Data
influenced
journeys**

profusion

Winning together

Our client partners



first direct



M&S BANK

C O T Y



**Data
influenced
journeys**

Profusion (UK office), 69-77, Telephone House, Paul St, London, EC2A 4NW

Profusion (Dubai Office), CNN Building No 2. Al Sufuoh Rd, Dubai Media City, PO BOX 72280