

# How to unlock the true value of data



## How to unlock the true value of data

Christian Annesley, guest contributor

Unlocking the value of data is no easy task, especially given that data comes in numerous formats. It often needs cleansing and deduplicating and there are restrictions on usage to consider.

In 2018, Gartner said the number of businesses enquiring about the challenge of [data sharing](#) had increased dramatically – by 180% in 2017 over 2016, followed by a 70% leap year-on-year in 2018. In their report *Implementing the data hub: architecture and technology choices*, Gartner analysts Ted Friedman and Andrew White discuss the challenge in terms of the need for a “data hub” that “delivers effective mediation of semantics, governance and efficient data-sharing across applications, processes, enterprises and ecosystems”.

### Get data flowing

Central to a hub architecture will be the technologies used to get [data flowing](#) into it from applications and other data sources, and then provisioning outward to consumers – internal just as much as external. These might include extract, transform and load (ETL) tools that support bulk or batch movement of data, data replication and data virtualisation.

---

They can also include app-integration middleware, such as the enterprise service bus, and message-oriented technologies that move data around in the form of message constructs.

Whatever tools are used, on-premise and cloud service versions are available to tap, and there are still other elements to consider, such as governance tools to help with data compliance and metadata management tools to tag and manage data flows better.

One of the big headaches for those tasked with developing a business's [data architecture is control](#). David Morris, a director of solutions for data integration specialist Tealium, says establishing ownership of the data that an enterprise handles, and getting it back under control, can be surprisingly hard work.

“Many corporates – and even plenty of relatively new SMEs [small and medium-sized enterprises] – have huge amounts of data, often sitting on more than one cloud platform like AWS [Amazon Web Services] or [Microsoft] Azure, but have also become too reliant on outside suppliers and their tools to maintain that data,” he says.

“Along that way, that means they have lost control of one of their most valuable commodities – the customer record. Getting back control is harder from this starting point, but very common. Ultimate ownership of the customer record matters if you want to build a data practice, and that means becoming less reliant on suppliers that store and manage data on proprietary systems.”

### Fragmented data

Morris says Tealium's customers suffer from [fragmented data](#) when they first get in touch, and this is compounded by the way data is tied up in others' systems.

"So the first step for many is to reclaim the data and get it back into an agnostic tool such as our Universal Data Hub," he says. "Without that, a company will remain hamstrung and limited in what it can do. An agnostic data layer unlocks things."

One of Tealium's customers, a large retailer with a sizeable distribution network, took a particular approach to its data challenge.

### Customer interactions

Morris says the retailer has a huge online presence, and constantly collects data from customer interactions. It wanted to provide a tool to suggest next-best actions, based on the intelligence it was gathering, that it could use to market to customers in real time. The data was put on an agnostic platform, and the tool was developed using Apache Kafka on AWS.

One tension often seen in organisations, says Morris, is the desire by data architects to use tools that smooth their daily lives and deliver clean and enriched data. But they still need flexibility to have full control of the data.

Once work on an overarching data architecture for the business is in place, individual data projects can start to flow and deliver commercial and customer advantage. That is the prize that companies need to hold on to as they develop their data strategy.

Couchbase Mobile is a collection of software comprising an embedded NoSQL database distributed with Android and iOS mobile applications and a middleware service to synchronise the locally stored data with Couchbase Server and other devices.

Newspaper headlines have often highlighted Ryanair's poor reputation for customer experience. The airline's mobile app has also attracted a number of specific customer complaints on ratings sites. To address these issues, Ryanair recently deployed Couchbase Mobile in a bid to increase its app performance and enhance users' experience.

Vladimir Atanosov, lead developer on the project, says: "Some apps don't travel well – and that was true of ours. Wi-Fi is not always readily available, and with a slow connection, apps can behave erratically. So we wanted to take a mobile-first approach and Couchbase Server plus Sync Gateway plus Couchbase Mobile suited us perfectly. It was easy to switch to because the native APIs [application programming interfaces] require almost no learning. It was also open source, with good technical support and a big developer community."

---

The result was an app that can easily manage semi-static data, did not need its own synchronisation and storage, and was far faster and better for end-users. In fact, Ryanair was able to overhaul its app performance without substantially re-architecting it by linking back to Couchbase running on AWS. “The booking process is at least 60% faster – cut from over five minutes to under two – and network traffic to drive bookings is 87% more efficient, cut from 80GB a day to 10GB,” says Atanosov.

Its users agree: ratings are much improved since the roll-out of the new app, supported on this fast-to-implement project by an underlying data architecture that made it all possible.

“Downstream activities need good data and good tools to deliver beneficial actions, but the data needs to be collected and stored in the right way in the first place,” he says.

By working on distinctive projects that use data in unusual ways, companies can also leapfrog rival businesses. “That’s another reason not to lock in with the suppliers’ systems,” says Stuart Mackintosh, founder and CEO of open source software data specialist OpusVL.

“You won’t gain market advantage performing standard data projects on widely used platforms. In the travel industry, for example, some companies have got ahead of others by investing some years ago in artificial intelligence and getting it working for them for enormous competitive advantage.”

According to the experts that Computer Weekly spoke to, data-led projects usually produce distinct short-term and long-term benefits. First, there is the immediate benefit that flows in six to nine months after having the data collected and owned by the organisation on its own terms, and applying it in some obvious use cases.

Then there is the benefit that will take longer to realise – to get into machine learning and iterate to develop well-tuned models over one to two years that can potentially deliver huge value down the line.

Experimenting is part of this journey – the more experience a company gains with its data, the more it gets out of it.

Industries where huge gains have been realised include publishing, where new-style personalised and subscription business models have been created by some to replace conventional sales models where margins were being eroded.

### **Financial services**

Another area that is seeing transformation through data is financial services, where incoming fintechs – challenger banks and the like – have created data-driven businesses on a different basis to the long-established incumbents.

There are many who describe data as the new gold rush, but the risk is that collecting data is not enough. Organisations need to have clear business goals – how will this data be used?

---

OpusVL's Mackintosh says any data-oriented project needs to begin with a thorough understanding of the project's primary business goal. "Companies need a handle on their pain points – their precise business needs," he says. "Technology is a tool to get them there. What must a digital system do if it is to deliver? If the aim is to automate certain functions, for example, you need to understand the data and the implications of automation well enough at the outset."

### Human process

This often begins with whiteboard diagrams that map out data flows. "It's a [human process](#)," says Mackintosh. "Working out the target usually means whiteboards and diagrams and imagination from those that best understand the business. A process like ETL is a one-off operation, remember. The end-goal is to facilitate a migration of data into the right format in the right places on the right platforms to make it sing."

Clearly, once a robust data architecture is in place, the data science team can work on the alchemy that turns raw data into the golden nuggets of business insight.

---

---

## Getting more CW+ exclusive content

As a CW+ member, you have access to TechTarget's entire portfolio of 140+ websites. CW+ access directs you to previously unavailable "platinum members-only resources" that are guaranteed to save you the time and effort of having to track such premium content down on your own, ultimately helping you to solve your toughest IT challenges more effectively—and faster—than ever before.

---

Take full advantage of your membership by visiting  
[www.computerweekly.com/eproducts](http://www.computerweekly.com/eproducts)

Images; stock.adobe.com

© 2019 TechTarget. No part of this publication may be transmitted or reproduced in any form or by any means without written permission from the publisher.