Legacy Infrastructure and Big Data

Sam Young
Analytics Development Leader

25th September 2018

nationalgrid
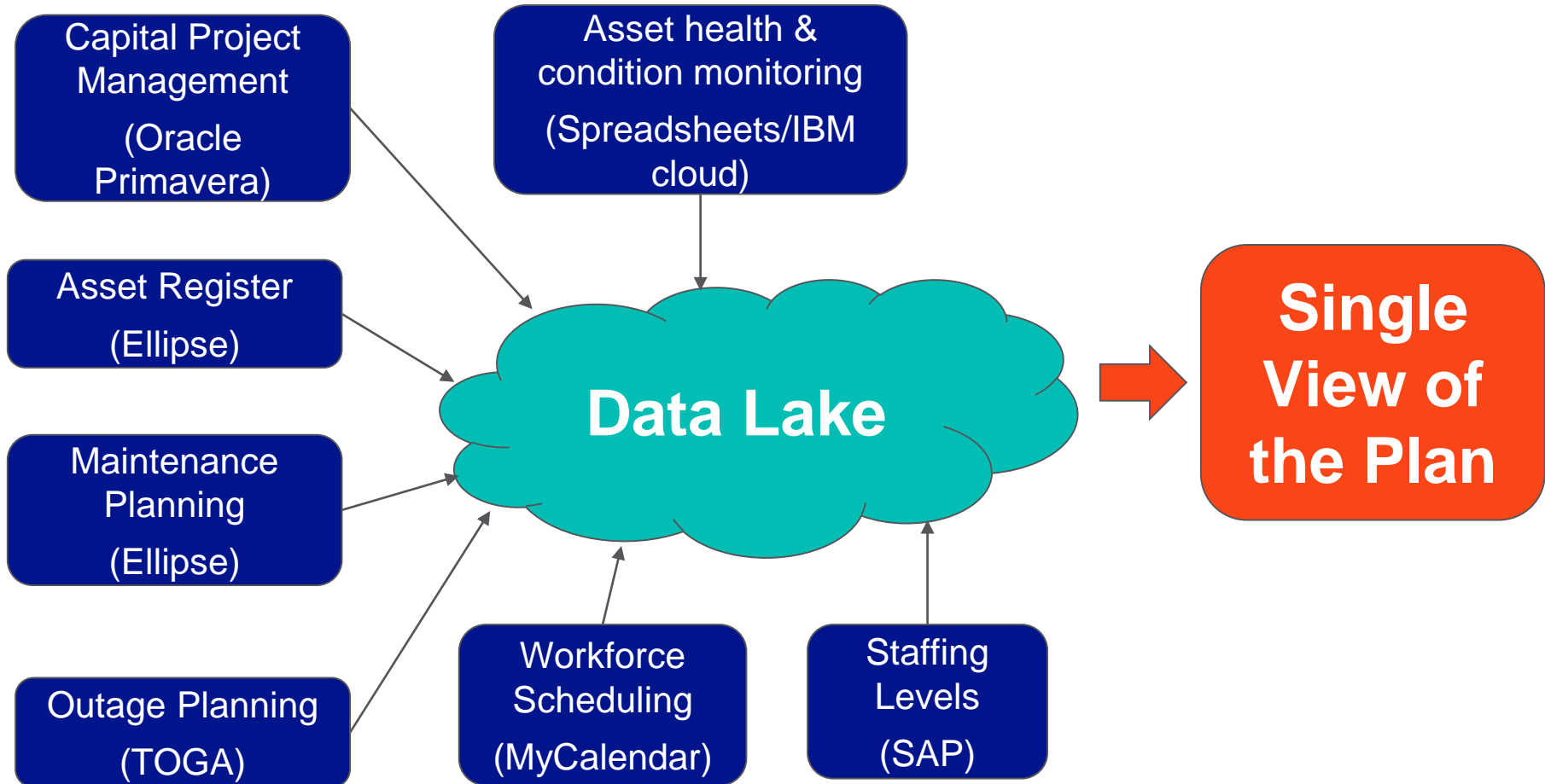
Contents page

The business need

Capital Project Management (Oracle Primavera)

Asset health & condition monitoring (Spreadsheets/IBM cloud)

Asset Register (Ellipse)

Maintenance Planning (Ellipse)

Outage Planning (TOGA)

Workforce Scheduling (MyCalendar)

Staffing Levels (SAP)

Data Lake

Single View of the Plan

The technology stack

Visualisation

Modelling

Integration

ETL

Storage

Performance challenges

We experienced a performance bottleneck on the ETL servers, which had an adverse effect on stakeholder engagement

We needed to continually re-optimise the ETL job schedule

Performance worsened non-linearly with new use cases

Predicting production performance was hard because pre-prod did not have continually refreshed data

ETL Schedule

On premises vs cloud

"It is difficult to make predictions, especially about the future"

Danish Proverb

On Premises

Expensive and slow to scale up infrastructure

Accurate initial predictions are therefore **very important**

Cloud

Cheap and quick to scale up infrastructure

Accurate initial predictions are therefore **less important**

Original:

On Premises

Revised:

Oracle Cloud

Data model

The main use case dominated the data model design

1. We have a large number of input tables from different systems

2. We have a clearly defined set of ways they will be used

Solution: combine all the tables from different sources into a handful of big tables with "everything" in

Implementing new use cases was made harder

Lack of intermediate tables led to rework/duplication and reduced performance

Original:

Big tables

Revised:

Intermediate tables

Data quality in legacy systems

Data quality in legacy systems always needs improving

Integrating systems reveals this poor data quality

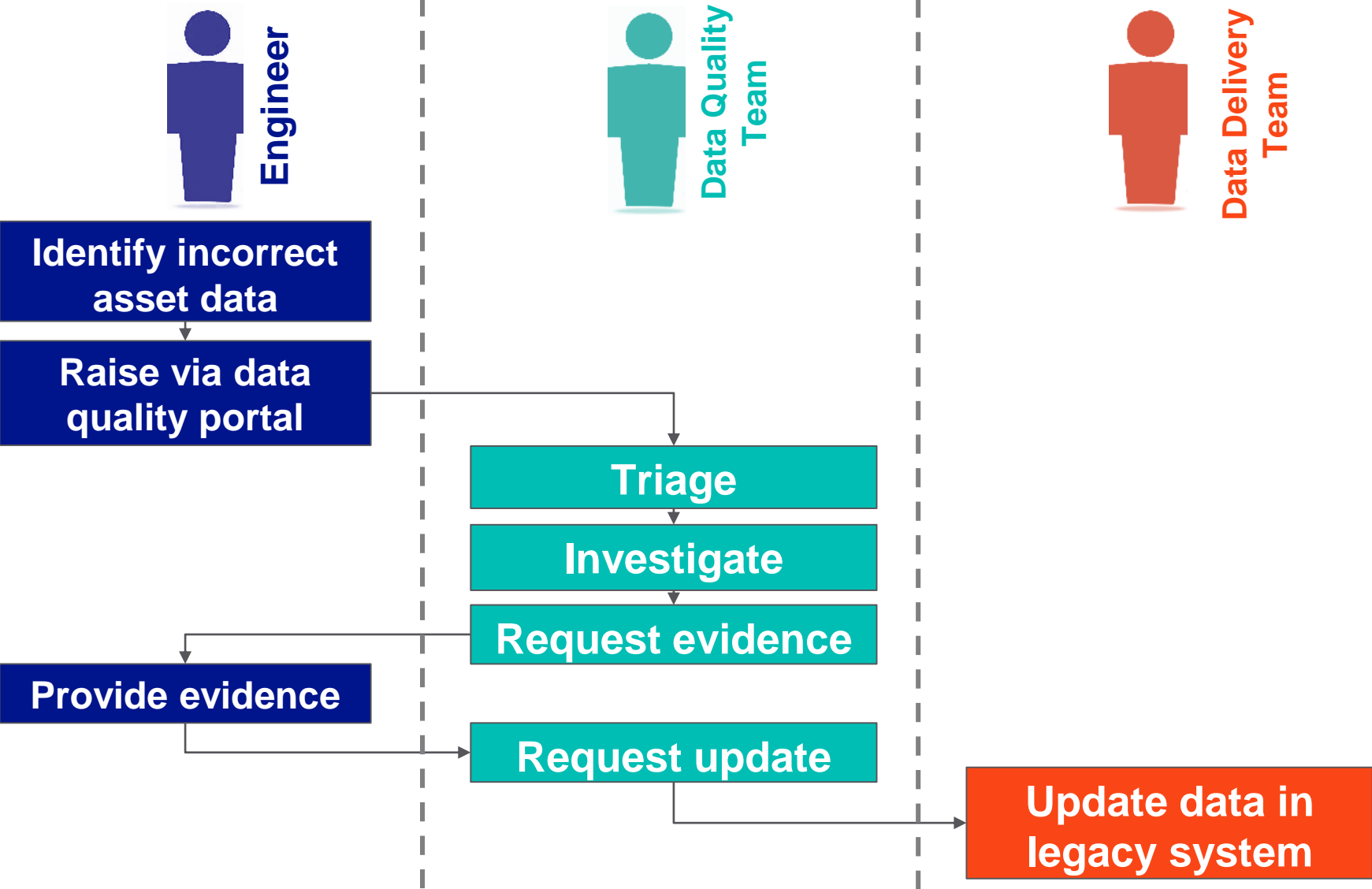New data platforms often offer better data management tools than legacy systems

These tools can be used to improve data quality in legacy systems

Business ownership of data can strengthen because data quality is more visible
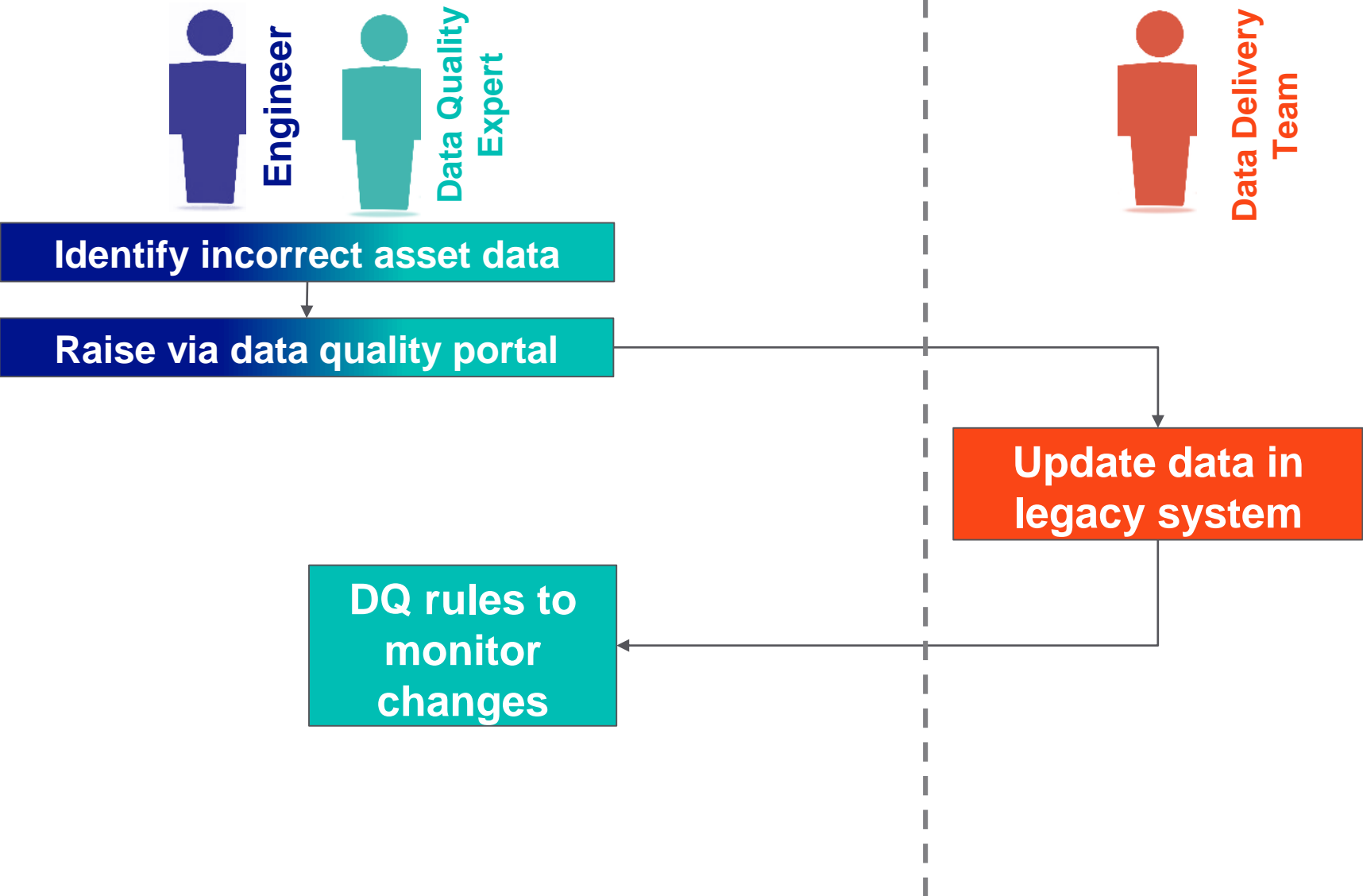
Example process for correcting data

Engineer

Data Quality Team

Data Delivery Team

Identify incorrect asset data

Raise via data quality portal

Triage

Investigate

Request evidence

Provide evidence

Request update

Update data in legacy system

Example process for correcting data

Engineer **Data Quality Expert**

Data Delivery Team

Identify incorrect asset data

Raise via data quality portal

Update data in legacy system

DQ rules to monitor changes

Agnes Allen's Law

> **"Almost anything is easier to get into than out of"**
>
> Agnes Allen

> **"Except for legacy systems, where it's easier to get things out than put them in"**
>
> Sam Young

Legacy systems are often not designed with nice APIs

e.g. raising new workorders

Robotic Process Automation may provide a workaround

Robotic Process Automation

Robotic Process Automation automates processes using existing front end interfaces to reduce the need for back end integration.

"Gaffer tape for systems integration" – not an enduring solution, but surprisingly useful

Ideal for well defined processes with limited permutations

Not robust to front end interface changes

National Grid has deployed it in various back office tasks

- e.g. translating shopping carts into orders in supplier system

We are considering the possibility of automating the creation of work orders in a legacy system using RPA
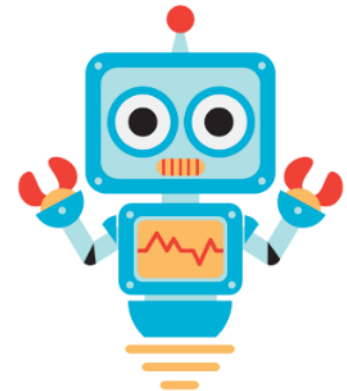
Conway's Law

"**Organisations which design systems are constrained to produce designs which are copies of the communication structures of these organisations**"

Melvin Conway

"Systems reflect the organisation that designed them"

e.g. siloed organisations produce siloed systems

Your systems and data structures are a mirror

Leverage insights into organisation

If you're struggling to integrate systems/data, there's usually an organisational reason as well as a technical one

The classic "these two systems don't have compatible keys" is a symptom of processes that don't join up well

We need to be pragmatic – implementing workarounds in systems is often necessary

However suppressing the symptoms is detrimental in the long term, so communicate the organisational insights and ensure they are followed up

Case study:

Tracking efficiency savings throughout the lifecycle of capital projects

Legacy systems fit a legacy organisation

Your current (and future) organisation is different from the one that designed the legacy systems

Legacy systems can make organisational change harder

Design integration to enable rather than restrict future change
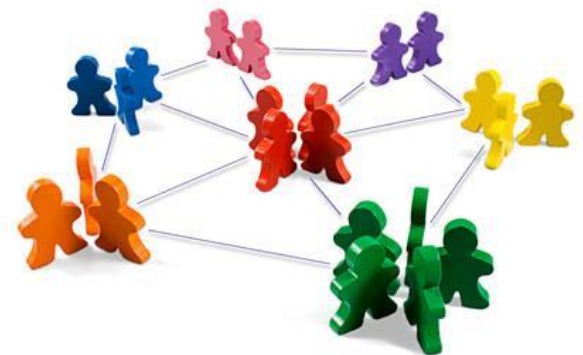
- e.g. intermediate tables vs big tables

Integrate relationships not just systems

One of the biggest benefits from major integration projects is the relationships that develop across the organisation

Work out ways to maintain and reinforce those relationships after the end of the project

- Coffee catchups

- Co-location

- Cross-functional teams

- Restructure

Conclusions

We integrate legacy systems to **enable change**.

So **value flexibility** and consider using "gaffer tape" for quick wins.

Remember that systems reflect organisations, so **help integrate the organisation** not just the systems.

nationalgrid