

The following presentation was given at:

SCAF Workshop

“Back to the Future – A Reappraisal of Cost Forecasting Techniques”

Thursday 8th February 2018

The Royal United Services Institute, London

Released for distribution by the Author

www.scaf.org.uk/library





BIG DATA – BEFORE YOU CAN ANALYSE YOU HAVE TO STANDARDISE

JAGUAR LAND ROVER - PRODUCT ENGINEERING FINANCE

Tim Brogan & Martina Sabova
08 February 2018

Confidential ©2018

Big Data

Topics Covered in Our Presentation Today..



- We will be covering some of the basic principles involved in Big Data, Data Mining, Analytics, and some of the possible implications and the legalities of doing it.
- We will not be looking into the mathematics and the potential tool choice when undertaking Data Analytics, but this could form the basis of a follow up presentation(s) at a later date....
- We will follow this with a brief overview of a current Data Analytics project in JLR

How Important is Data to Business?



“Data is the new oil.”

Ajay Banga
President and CEO of Mastercard

“Big Data is Neutral: A tool for
both Good and Evil”

Rick Smolan
Co-author “The Human Face of Big Data”



The Golden Rule

Before You do Any Big Data Analytics.....You Should Define
the Question You Want to Answer...

We are at the dawning of the Big Data Era

Extracting it will become more difficult and expensive.....



- Big Data is the new “Buzz Word” everyone wants in....
- More and more jobs appearing with Data and Data Analytics in the title...
- Opportunists are extracting the easily accessible data near the surface...
- The Wagon trains of Analytics Developers and Data Scientists are on the move.....



We are at the dawning of the Data Era

Extracting it will become more difficult and expensive.....



- As this easily found data dries up..... Good data becomes more difficult to extract.....
- How much drilling and waste extraction will need to be done to get to it?
- The “old” tools have been good up to now.... But can no longer be relied upon to deliver in the new era



The Good Data

“Big Data- How do I know if I have got some?”

Characteristics -The 5 V’s.....



- **Volume-** The quantity of data generated and stored for analysis, Big Data does not sample.
- **Variety-** The type of data you have to analyse, text, images, fills in missing pieces.
- **Velocity-** The speed that the data is generated.
- **Variability-** Inconsistencies in the data set, can slow down the analysis
- **Veracity-** The data quality that is available can vary enormously

“Big Data - Where Do I Store it All?”

The Data Lake.....



- Contains a variety of data and data formats
- The Data in the lake should always be in its natural format
- Analytics is the primary reason for the growth of data lake architectures
- The challenge is not in creating a data lake.. It is keeping it current and relevant
- Beware!, there are no rules to say what defines a data lake
- Don't let the lake become a swamp..... the data loses its inherent value to the business

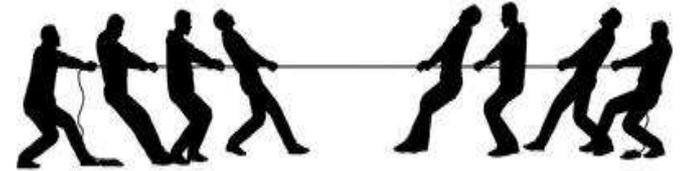


Who in the Business Owns All of This Data...?

Big Data is a Corporate Asset



- **Option 1** – Nobody knows.. No strategy or policy in place... a free for all.
- **Option 2** – Owned and managed by Corporate IT, however, the people in IT rarely understand the true nature of the data they are being asked to manage, they can be easily drawn into being more concerned about the technology than the data
- **Option 3** – Giving corporate responsibility to a single function e.g. Finance, usually meaning that the function can be quick to respond to data requests
- **Option 4** – A matrix approach between IT and a corporate functional team, probably the best approach as it draws on the strengths of both elements, creates a challenge in organisational complexity and in ensuring that both the technology and the data are available and at the latest standards.



Is Big Data for Everyone...?

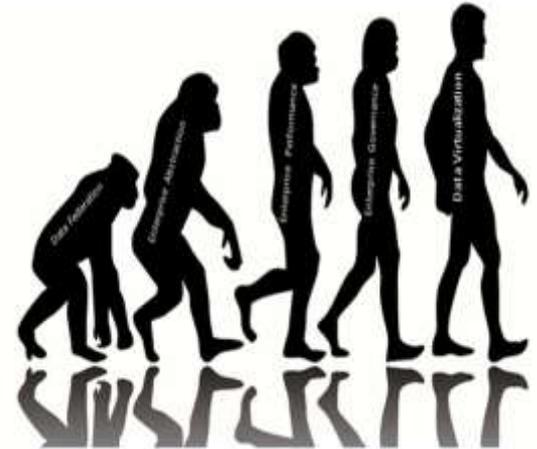
Are You Really Ready to do Data Analytics?



- All companies will be at different starting points in their data evolution
- If you only use your data to monitor your business performance and to do the accounts, you are unlikely to be taking any sort of lead in Big Data and Data Analytics
- If your company has one function that is totally dominant is analytics you are unlikely to be able to take the matrix approach very easily
- How willing is your company to invest in bringing in, or growing the capability..?
- Does your company have a plan to gather and analyse your data...?



- **First Steps / Immature** – No central reporting, Departments create their own reports, use of shadow IT, Excel and Access seen as the tools of choice, no enterprise level data sharing
- **Evolving** – Some level of Enterprise data, still has some overlaps, silos still exist in data strategy, competing versions of the truth.
- **Maturing** - Central Data Repository in place, common reporting sets, corporate agreed data definitions, Data Analytics and Forecasting possible and can help to set strategy, but not 100% reliable, data is an enabler, but not fully integrated into corporate strategy.
- **Differentiating** – Company uses Business Intelligence based decision making to shape its strategy. Data can be pulled from across the business to answer any question, and to set direction for the future.



Big Data

Big Judgement.....



- Collected data, big or small will only really inform you about what has happened or at best what is currently happening.....
- If the future is likely to be different to today... no amount of analysis of your data will be able to accurately predict it.
- To predict a different future.. You need to thoroughly understand the system dynamics in play, this can only be achieved by taking a theoretical approach
- A combined approach of Big Data Analytics and systems modelling could be a better approach



- While Data Mining itself is perfectly legal to do, the data you mine and the patterns you may find within it may have implications, in its legality, the ethics of doing it, and invasion of personal privacy.
- Different countries will have different laws on these matters
- Even “anonymised” data sets could potentially have enough information to allow for individuals to be recognised
- Therefore the following considerations should be addressed:
 - The purpose of the data collection, and any data mining envisaged on it
 - How the collected data will be used
 - Who will be allowed to mine the collected data and use the findings
 - The security protocols under which the mined data will be stored
 - The updating process



Big Data

Data Ownership



- Data has become a highly valuable asset within any company, but definitions of data ownership and mechanisms for claiming and protecting rights of ownership are anything but clear.
- There is currently no European Legislation that specifically regulates the ownership of data. Other legislation such as Copyright, Trade Secrets etc have to be invoked to cover for this. But legislation such as Trade Secrets were not set up to cover data and so can be regarded as potentially not being fit for the purpose.
- Data ownership has been defined as the right to control the data and to claim the profits generated from it. Controlling the data and who has access to it is crucial to ensure that it is not copied or distributed more widely the absolutely necessary.
- Care should be taken when working with any 3rd parties where access to data is possible that ownership of the data is very clearly agreed to vest with your business and that ownership is not passed to the 3rd party.





JLR Case Study

The “Final Status Review” Database

Product Creation and Delivery System Overview



- The Jaguar Land Rover process used during the development of a new vehicle is known as the Product Creation and Delivery System (PCDS)
- There are nine specific elements within it:
 - Project Governance
 - Product Definition
 - Design
 - Integration
 - Commodities
 - Software
 - Prototypes
 - Investment
 - Launch



The Final Status Review Gateway Process Overview



- Re-assess the Strategic Intent
- Review the costs with other stakeholders
- Agree the data freeze point



The Final Status Review Database Project Purpose

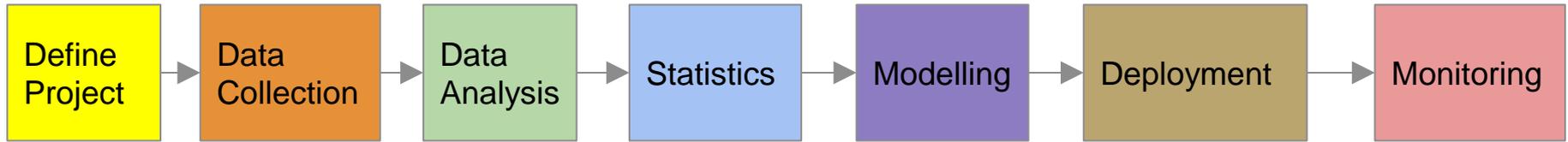


- We were asked if we could create a central repository for all of this data in a consistent data structure and at a number of levels of indented cost structure.



Big Data

The Analytics Process



Final Status Review

Data Storage



- Historically Product Engineering Finance have held the frozen data in a series of non-linked Excel and SQL files.
- These files have been called a “Redbook” ever since JLR became part of the Ford Premier Automotive Group, and have continued to be called this in JLR ever since.
- The Redbook contains both the investment cost of the project and the recurring part cost of the vehicles (as an each cost per variant)

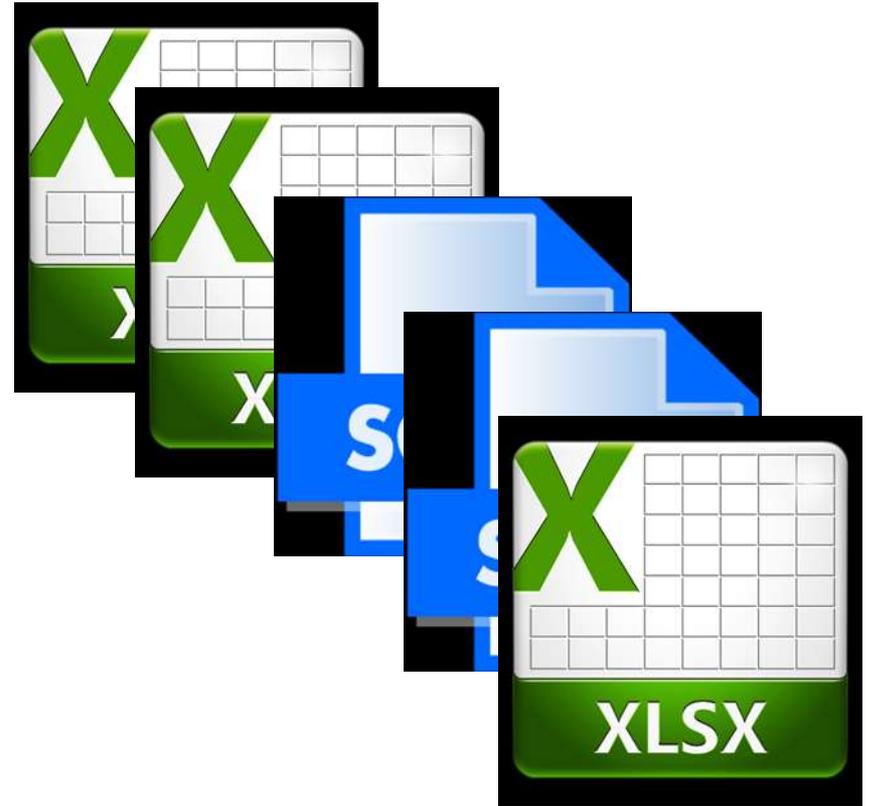


Final Status Review

What we Found When we Started to Analyse the Data - Issues



- Data Formatting
- Uncontrolled Sources
- Data Quality
- Data Definitions
- Multiple Versions
- What is important?



Final Status Review

How we Approached Finding a Solution to the Issues we Found



- Standardise
- Profile
- Identify Users and stakeholders
- Determine data quality
- Process & Practices
- Data Governance



- Get the Data Out-there! - initial data sharing with colleagues
- Look for potential solutions
- Evaluate each solution
- Finalise the solution
- Load data to the system



And Finally.....

Our Conclusions on Big Data



- Data is all around us, all of the time
- Data is expanding at a tremendous rate
- The relationships you may think you have found in a data set... may not be..... unless you can prove causality
- You cannot get around good practices in how you gather and store data. If you mine poor data... you will get poor or incorrect answers
- Data Analytics are not a substitute for rational human thought and observation....





THANK YOU

ANY QUESTIONS.....?

(Tim Brogan)
(Data Services Manager)

Jaguar Land Rover
W/1/26 Abbey Road, Whitley
Coventry CV3 4LF, UK
jaguarlandrover.com

(Martina Sabova)
(Senior Data Analyst)