

# Data – ‘The Good Oil’

It has often been said that data is the oil of the new age<sup>1</sup>. Data is an enormously untapped reserve in our modern business world, and just like oil, is immensely valuable. However, just because you have the oil does not necessarily mean you have anything useful at all, it only comes into its own when you have a machine to put it through or you make something from it. Oil has been used as a lubricant and a fuel for almost two centuries. It is the basic building block of many of our modern resources such as plastics and roads. We are only now beginning to see the massive potential of the new oil.

Today as we strive to use that oil (data), in new and innovative ways to improve the human condition, we have come to the realisation that it is well, not always cracked up to what it seems to be. To drive the analogy with oil even further, our data is “dirty,” and so must be “cleaned” to make it useful. We have all heard this before, yes, real world data is duplicated, inconsistent, incomplete, full of errors, (either intentional or unintentional) and with the sheer volume of data that is being produced daily this problem will only become more amplified. To run a machine on oil at peak performance we use an oil filter. The importance of an oil filter is to remove impurities and small particles that can damage the engine. The same is true with our data.

As we drive many critical decisions across many different industries, (some of which can greatly impact our health and wellbeing), it is particularly important to filter out the impurities of that data, to make it clean, or in other words, trustworthy. Now trust is a particularly important consideration in today’s world, we need to make sure that any decisions being made, either by corporate executives, government or big tech are based on reliable, trustworthy data. We often find that data analysts will go elsewhere to find their data sources if they do not trust the data, and this can lead to many frustrating, confusing and sometimes dangerous scenarios.

As we begin to develop more sophisticated machines for processing data, i.e., quantum computing, and applying Machine Learning to develop Artificial Intelligence platforms, it has become more important than ever to consider the Single Point of Truth, or SPOT, as the foundation of our analytic solutions. We have had a great head start on this over the past 30-40 years with the help of ideas from Ralph Kimball<sup>2</sup> and Bill Inmon<sup>3</sup>, (the technological differences between these two giants of the data world are still hotly debated). We have seen recently however that analysing the likes of big data and corporate information data can be extremely challenging as the Data Warehouse is not as

flexible as we would like it to be, so today we bring together the best of both worlds, the Data Lakehouse<sup>4</sup>.

Here at Fujitsu Data & AI we pride ourselves in the fact that we are a world leader in Data Analytics and AI, providing services to unlock the potential of data across diverse sectors using Machine Learning and Databricks, which is typically built on the foundation of a Data Lakehouse.

If your business needs help with their data, please contact a Fujitsu Data & AI specialist now.

**Contact**

Fujitsu Data & AI  
+61 3 9924 3000

© Fujitsu 2022. All rights reserved. Fujitsu and Fujitsu logo are trademarks of Fujitsu Limited registered in many jurisdictions worldwide. Other product, service and company names mentioned herein may be trademarks of Fujitsu or other companies. This document is current as of the initial date of publication and subject to be changed by Fujitsu without notice. This material is provided for information purposes only and Fujitsu assumes no liability related to its use.