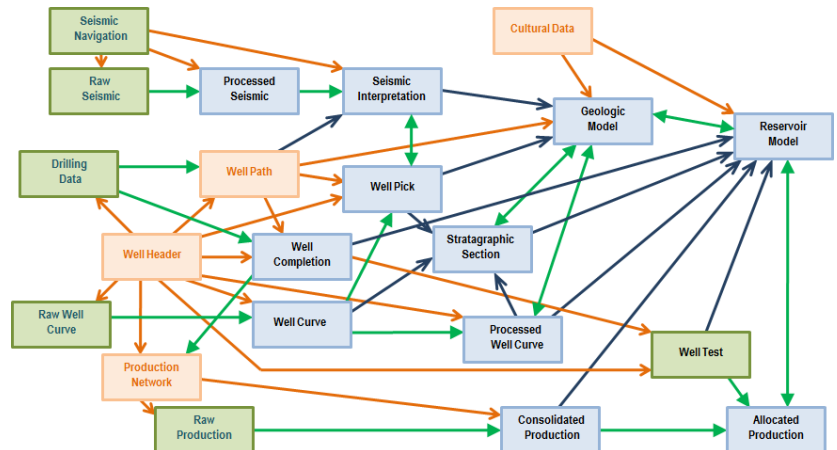


Which data is “managed”?

Oil companies rely on many different sets of data to be effective. From the locations of shot points on a 2d seismic line, through to an asset model that combines dynamic subsurface simulation with surface network modelling. This glut of data groups interact in complex ways, key values are defined by one expert’s integration of figures which in turn were someone else’s estimate based on something else.

In the past¹ I’ve illustrated this with a picture of various broad “data categories”. Of course, this image grossly simplifies the actual level of interactions in even the smallest oil company, but even this version illustrates a key point and raises an obvious question.



The point is that the majority of data is the result of some kind of process performed by an expert. It does not arrive in a “raw” form, as a direct measurement of some physical thing, but rather is some clever (and experienced) person’s “best estimate”.

Consider the “neutron” well log, usually the one being used will be the result of a petrophysicist depth shifting and merging a number of runs, removing the obvious spikes, scaling and tidying up the curve in other ways. Now of course this result does embody actual magnitudes of the neutrons scattered, but the primary goal is not to slavishly follow the original (inevitably flawed) measurements but rather to be a valuable aid in estimating the rock’s actual properties.

In this diagram there is a clear difference between the data categories on the left side and those on the right. The data on the right can visibly be tied to business decisions, the planned path of a new well, for example, might be most noticeably influenced by the structural and dynamic models. This means that we can, in principle, work out the business value of having that data (add up the value of all the decisions that depend on those items and divide the sum between the things that influence the choices made). However, while we can estimate how much value they generate it is much harder to work out how much they cost because they are the result of investing an unclear amount of professional’s time integrating insights from a whole range of different sources. In contrast it is much easier to estimate the cost of the data on the left. The final neutron curve mentioned earlier relied solely on the logging contractor’s measurements and the petrophysicist’s time, but understanding which future business decisions that curve will impact is much more difficult.

One would hope that the value to the company increases as more expert’s time is invested, if a structural model is less valuable than the seismic interpretation its based on why spend the effort to create it? At the moment the activity of “Data Management” in most oil companies only deals with the data towards the left side, the seismic field data, the log curves and so on. But shouldn’t the more valuable “results” also be properly managed and stored for posterity?

¹ e.g. Fig 11 in the paper “Applying DAMA to Oil Industry Data” available at <http://dm4ep.com/data/shawtin-slb-pnec14-a4.pdf>