

Can I get some insights, please?

Over the years, I have come to somewhat dislike the term 'insights' almost to the same level as, say, a 'Data Lake'. And that's saying something. Not because these concepts themselves are related that much (they are to *some* extent, of course). But, because to me personally, they both conjure the same feeling: a mixture of annoyance and desperation. One of the reasons is that since the word 'insights' is part of my job title, there is a general view that insights are something that you can just 'mine' from 'big data' and are there to be 'delivered'.

My experience however has been that insights are generated from working closely with the business over a longer period of time, and with a dedicated focus of honing in on explanations that are relevant to meet a *specific* business goal. Insights do not appear from data, and 'shaking' data vigorously does not make insights fall out of it. Insights supplement data analysis by proving or disproving *detailed* hypotheses or assumptions that the initial analysis may have brought to the attention. Furthermore, insights are sharpened by implementing a feedback loop that continuously addresses gaps in knowledge and information over a longer period of time and across various areas of the business. This article aims to explain how we can achieve this cross-functional collaboration, which I believe is a requirement to get 'insights'.

The way I usually introduce this topic is by emphasising the 'enabling' of all of this: the data itself and the hard work required to bring data together, and address any gaps. Also, do not underestimate the effort and diligence required to ensure data is of reasonable quality.

Ownership of this collection process lies in the realm of Data Warehousing, Data Integration and to some extent Data Governance. They can be seen as the custodians of the data, which by the way does not mean that ownership of the data lies here as well. Ultimately the business owns the data (or should take ownership if they haven't done so already) and should determine its fitness-for-purpose, as they will be the ones that ultimately will reap the rewards of the data.

To support the above, as an Insights Manager, my intent is to mold this data into 'data (Lego) blocks': defined and auditable data elements of sufficient quality, with which you can build many things.

The sky is the limit, or until you run out of blocks.

The idea is that exposing these 'data blocks' provides basic information to Subject Matter Experts (SMEs), which hopefully sparks creativity and subsequent deeper analysis. The ideas that emerge from this creativity are the nutrients to cultivate 'insights'. Self-service (managed self-service, that is) plays a role here, too.

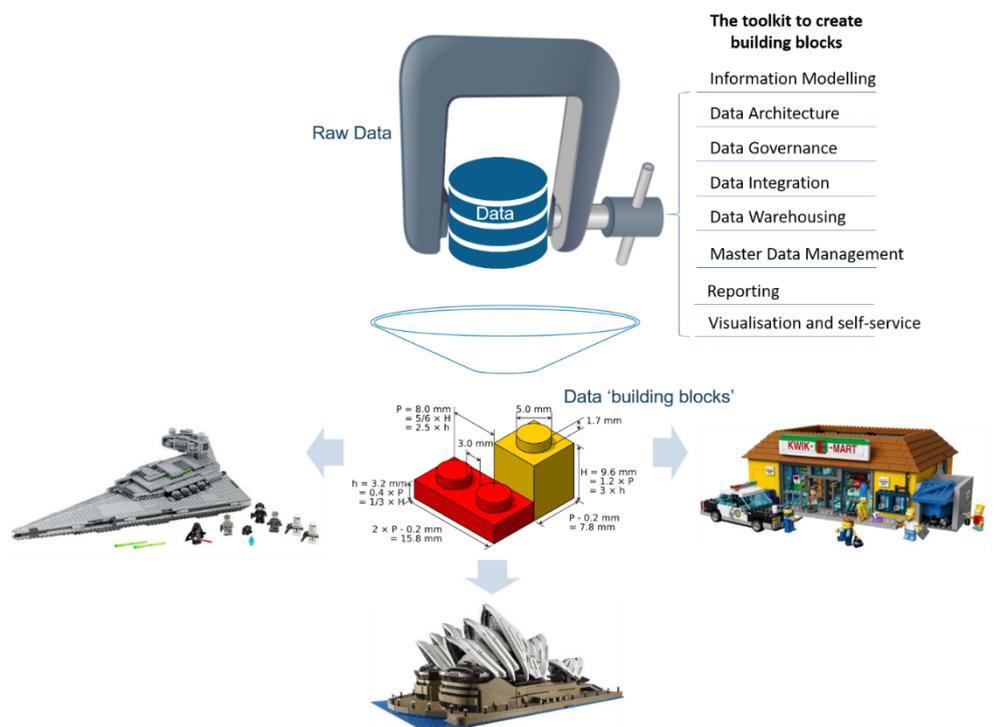


Figure 1 - with the right reusable data blocks you can build anything. Data availability sparks creativity.

The Book With All The Answers

We have a book at home called 'The Book With All The Answers', which is absolutely great if you're looking for *any* answer to *any* question. It is basically a book full of answers such as 'yes, you should totally go for it' or 'that would be the biggest mistake of your life'. Every page has a specific answer printed and all you have to do is say the question out loud, open the book without looking and you have your answer. There are times that this is used rather heavily (you know who you are!).

There is a work equivalent of this as well. Invariably, I am asked to 'give some insights', or 'can I get some insights on our XYZ?'. We sometimes (jokingly) respond by saying 'give me a minute and I'll run the Report With All The Answers' or open up the 'Silver Bullet View'. Because it doesn't work like that.

Any kind of analysis *always* starts with a business problem. Or better yet: any kind of analysis *should* always start with a *clearly defined* business problem.

So, how does it work? While working on analytics and insights can start at any time and with incomplete data (as long as there is a clearly defined problem), I believe it should be linked to a holistic approach that focuses on incrementally delivering more information to support the insights process. An approach that focuses on addressing gaps in data to sharpen the analysis over time. From this perspective, the continuous delivery of insights is part of a bigger process that is supported by data availability: the available set of data blocks.

It is also about finding a way to engage with the business on the 'what's next?', 'so what?' or some other specific topic. These would typically be use-case driven: aimed at achieving an actionable goal, but can also involve investigations for specific topics to understand why things happened.

Another way of looking at this is that fundamental, more basic information needs ought to be satisfied first before you can talk about generating insights. Insights require cultivation, to grow over time and don't magically appear out of thin air. You can picture this as a *hierarchy of data needs*, in a similar way to how human and / or organisational drives are described in Maslow's pyramid of needs.

Maslow's pyramid reinvented

You probably are aware of Maslow's Hierarchy of Needs, a model of human longings which is typically represented as a pyramid. It can be used to understand what drives (motivates) people, and that more basic (lower) needs require to be fulfilled first before the 'next level up' becomes relevant. For example, a basic need would be food, water or rest. These fundamental needs are required to be met before relationships and accomplishments become focus of attention. And so on.

This approach can also be applied to data needs, which in my view culminate into 'insights' at the top of the pyramid. This is the classical 'Data, Information, Knowledge, Wisdom', or DIKW pyramid. The origins and source are disputed, but generally Russel Ackoff is credited as being the first to broadly cite this concept in 1988.

In any case, I like this way of thinking and have been using it to positioning insights as well as capabilities such as Data Warehousing, Business Intelligence and Advanced Analytics (Data Science) in general. The 'pyramid of data needs' provides a good overview of what kind of requirements and capabilities a business should have in place in order to support ongoing insights generation.

I took the liberty of adding two additional layers to the original hierarchy of Data, Information, Knowledge and Wisdom: an explanation with examples in the middle column and ownership topics on the left column. The result is added below.

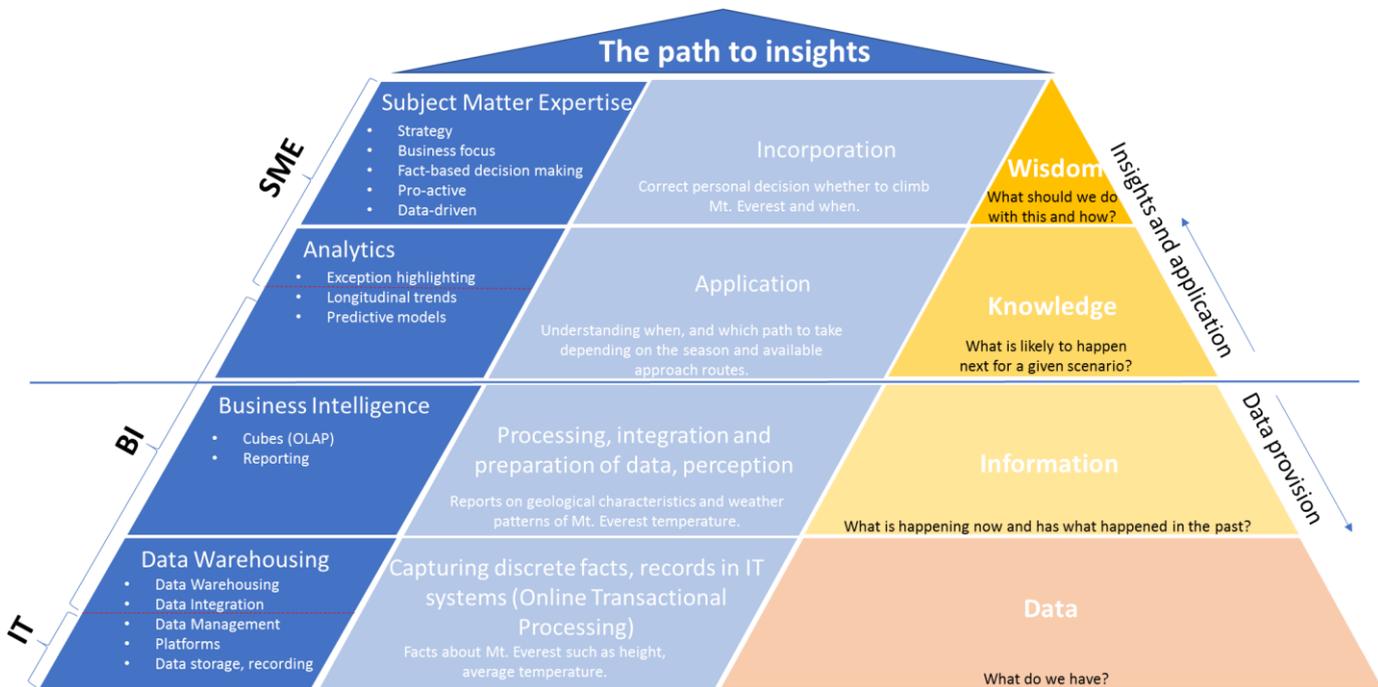


Figure 2 - The Pyramid of Data Needs

Adequate recording of events, transactions and descriptive *data* is the fundamental requirement. Without data, there is very little to report or analyse so we can define this as the most basic data need. Data is captured by operational (Online Transactional Processing) systems, which are typically managed by the IT departments through a combination of infrastructure and application management. Data is subsequently 'sourced' from these transactional systems into Data Warehouse (DWH) solutions through some form of Data Integration or emerging alternative techniques such as data virtualisation. Please note that not all data is available in neatly organised environments such as databases. More often than not there is a significant amount of (potentially) relevant data available in semi- or unstructured formats as well, which is something to bear in mind. Regardless of how this is stored though, it needs to be recorded *somewhere*.

Management of these solutions is usually part of the mandate of the Business Intelligence (BI) team, although in some organisations the responsibility of capturing (DWH) and reporting (BI) are separated (not my personal preference, though).

This answers the question about what data is available in principle, and can also be used to address gaps in data collection. Modelling plays a part here, usually through a conceptual or logical information model to highlight the ideal and as-is states.

Once the fundamental need for data to be recorded is satisfied, and data can be accessed in a fairly consistent way including access to previous states of the data (the history of changes), various questions can be answered such as 'what is happening now?', and 'what happened in the past?'. This is one of the core domains of BI and provides an important enabling service for any organisation. Providing consistent access to information also helps guide the organisation towards conformity of definitions and meaning, another key BI responsibility. Limited extrapolation or analysis is done 'forward-looking' in time, but the majority of focus is the 'rear-view mirror'. The ability to achieve this is another major requirement to proceed up the pyramid: raw data is processed into *information*.

There is a strong relationship between BI, DWH and Analytics although analytics in some forms can (and do) exist without the former. I believe however that structural application of analytics requires a more holistic approach where both information and (business) knowledge 'gaps' (uncovered through the analytics process) are addressed by the BI and DWH areas on an ongoing basis. A BI feedback loop if you like. In the same vein, validated findings and completed analysis can be incorporated as Business-As-Usual by the BI team. Data and business knowledge gaps go hand-in-hand here, as a lack of accessible information is likely to be a major cause for not fully understanding a particular area of the business.

In any case, Analytics (Advanced Analytics) incorporates various approaches for using statistics / modelling to predict, explain or explore relationships and causality in data. Predictive modelling applies a (carefully selected) statistical model or data mining algorithm to data for the purpose of predicting new or future observations. This topic covers, as an example, propensity analysis and generally looking at trends from a more statistical view such as calculating seasonal influences. Practically speaking Analytics is a way to help answer questions such as 'what is likely to happen next?'. In this context, it is interesting to consider that predicting something successfully using an analytics model does not equate to (fully) understanding why this behaviour manifests itself.

At this level, an organisation would be able to develop *knowledge*: an understanding of a specific topic including weighting different approaches. Each 'approach' is the result of a specific analysis for a specific defined topic or question. But how are these topics or questions selected?

We have now approached the core of this article, how to connect and engage with the business to achieve insights. While obviously relevant for BI as well, it is especially important for Analytics teams to work closely with the business. Another way of saying this is that (business) questions need to be very specific in the way they are formulated. It is for this level of detail that Subject Matter Experts(SMEs) come in, as true SMEs understand these nuances and generally have a clear idea on what the intended outcome needs to be, or what is generally the expected result.

If the outcome of the test is different to what is hypothesised by the SME, then you have your first insight.

SMEs need time and focus to work with the BI and Analytics teams. But, they also need to connect with the strategy and decision making side of the business. A great [post](#) by Tom Breur explains this really well. With this process established, the foundation is ready for the last level of organisational data needs. At this stage, a specific direction of the business can be supported by the required data in a processed and accessible way and various tests have been executed. Obtaining *wisdom*, by incorporation of information, is where insights live, because this is where various options have been defined, weighted, understood and ultimately used in a decision-making process that aligns with a specific strategy and intent.

These mechanisms need to be considered when working towards, or organising, data-driven initiatives. These initiatives need time, focus, cross-functional collaboration (e.g. SME and information specialists) and perseverance. Ultimately it comes down to understanding intent (strategy) and the ability to define the metrics relevant to the *specific area* of expertise and at a very detailed level, be it direct marketing, contact centres or risk modelling. The BI, DWH and Analytics teams can then work with the business to make this a reality. This means that delivering insights is a team effort across the organisation and requires time, focus and collaboration with the various areas to provide the data requirements following the 'data pyramid of needs'.

More insights

As stated before: every analysis should always start with a clearly defined business problem. The pyramid of data needs aims to explain what the requirements are to enable this, culminating in collaboration between the business (SME) and information specialists. If you're not an SME and are tempted into accepting delivering insights in areas that are not clearly defined, then your first problem is to figure out what the problem is to begin with. Hopefully this paper helps to avoid this pitfall.

This post admittedly has a very bottom-up view about working with data and turning it into insights. But my experience is that without these foundations in place the analytics process can be very slow, because gaps in data availability will be encountered repeatedly. Ironically, this is almost counter-intuitive to the 'big data' idea that we're swamped in data. By incorporating a data management approach or platform into your strategy, where the 'data Lego blocks' are being added to on a use-case basis you will be able to support the ongoing analytics process and bridge gaps in data and knowledge over time. Having said that, this is ultimately dependent on a combination of factors like the topic, the scale and the way data is managed for a particular organisation or subject area so your mileage may vary.

Market research (both quantitative and qualitative), is also a factor which is not covered in this paper. This provides another perspective on generating insights (which still requires a fairly well-defined topic though). In all cases being mindful of the types of skills, the specific focus and data needs will help putting together the right team for the right results.

At the end of the day, generating insights is an iterative cycle, a two-way street, that starts with strategy (focus) and gradually informs and improves understanding for a specific topic if you have the perseverance to stick with it. Another way of looking at this is that just having strategy and information is not enough. What we should work towards to is to create the feedback loop between the two areas and continuously improve this over time.

The sheep with five legs

This focus on cross-functional collaboration is opposed to the notion that an analytics specialist needs to be able to be an expert in both the business and statistics: a 'sheep with five legs'.

I'm not really sure this translates well from Dutch (my native language) into English, but a sheep with five legs in Dutch is basically a super talent: someone who is above-and-beyond versatile and can act as genuine catalyst or intermediary. The employee everyone is looking for. In the context of this article our 'five-legged-sheep' would be an advanced analytics specialist who also happens to be an expert in one or more areas of the business. Someone who, for instance, truly understands marketing, aviation, fast-moving-consumer-goods - whatever is relevant for the business - and at a practical level.

I don't disagree that finding someone that can combine a business specialisation with analytics expertise is a great asset. The point I am trying to make is that, especially in a large organisation, a single person or small team cannot be 'the SME' across all areas of the business. Instead, what you may want to hire for are so-called generalised specialists: people who excel in some area, yet are conversant enough in multiple adjacent domains to be versatile.

You therefore have to consider organising the teams to either focus on dedicated areas and hire for this purpose, or focus on broader capability and engage with SMEs in the business. Or maybe launch your own five-legged-sheep breeding program?