baufest

evolving business together

7 Steps

to Creating a Data-Centered Business

Let's say you've heard about this magic thing called data.

You've read the news, posts and ads, all claiming that data is the competitive advantage of the 21st century (which it is) and that you can tap into it's potential (which you can) to make your make better decisions (which you will). **You really want to get into this data thing, but... Can it be done? How? And most importantly: Where should you start?**

While there are no easy answers, there isn't one that is impossible to put into practice either. Depending on how far your organization has moved along the data path, the answers will vary on a case-to-case basis, so let's try to take a look at the most common milestones where this question pops up.

▷ ▷ ▸ ▸ ▸ ▷ ▸ ▷ ▷   ▸ ▸   ▷ ▷ ▷ ▸ ▷ ▷   ▸

01.
STEP ONE:

I have nothing...

The first exercise is to ask yourself: what do I wish to know before and after I make a business decision? This is the time to be visceral. Think of all the business decisions you have made before and the ones you **have** to make, and list all the information you wish you had.

If you **don't have anything**, don't try to jump into indexes, ratios and correlations from the start. That comes later. At this early stage, what is important is to define the relation between information and decisions in a concise way: "I wish to **know** how my customers feel about the products I sell so that I can decide what part of the catalog I should expand", "I wish to know if my employees struggle to meet their goals so I can decide whether to hire more people or not".

Congratulations! Now you have a clear driver for your data endeavors. Now let's see how to enable it.

02.
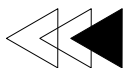
I know what I want to know. Now what?

Now that we know what we want to know, but don't yet know how to get it or read it. Now is the time to think of what number will give me that information in a quantified way. Now is the time to think of the answers to our questions in terms of METRICS. Now is the time to think of ratios, indexes, aggregations and such.

What information represents customer happiness with a product? What number will best indicate an employee struggle to meet a goal? And most importantly, **what data will I need to build those numbers?**

03. STEP THREE:

I know what I want to know. Now what?
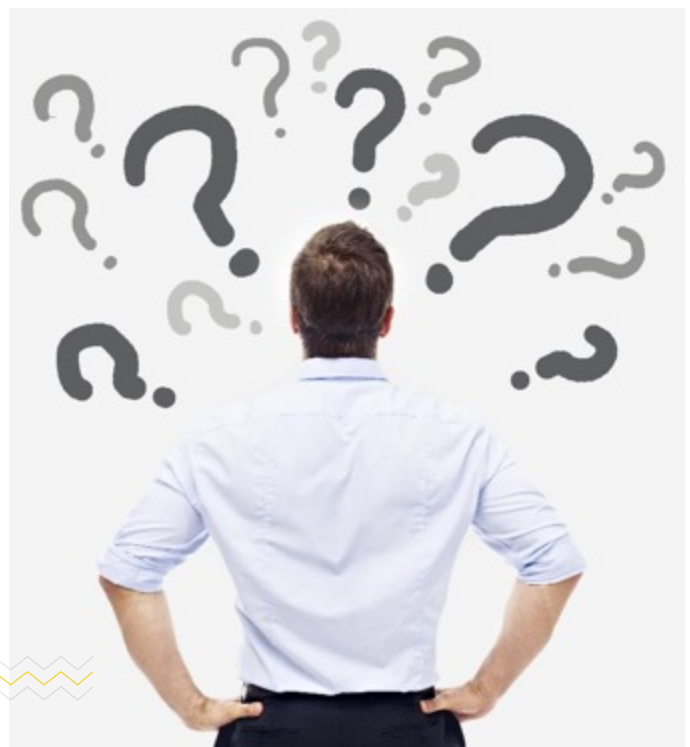
Depending on an organizations size, growth speed and antiquity, there is a big chance that there are a lot of systems lying around, generating unused yet exploitable data. So, once we know our metrics, it's a good time to build a data catalog. This means defining all the business terms involved in the metrics, avoiding IT specific terms, and surveying the existing systems to understand if the term is represented in some capacity in those systems, and if it adheres to the business definition we provided. These are the building blocks of our metrics.

This is also the best time to define the level of quality we expect from that information in order to trust it.

What data is acceptable to be missing? What is the reasonable threshold for the aggregate of any given value? How old can a record be before it's obsolete?

Some of the metrics you are looking for, will already be lying around somewhere among the tables and APIs your organization already has, and it will only be a matter of putting 2 and 2 together. Some will need a bit of remediation before they are fit for use. Some of them will throw new questions as a consequence, and these will be drivers in their own right. Some might be informally written in pieces of paper across the organization, and you will need to find a way to structure it in a meaningful way. Some will require the organization to rigorously define business terms that might have been ambiguous all this time without representing a problem. Some will demand a company wide effort to create the data points needed in order to reach it.

Regardless of where the answer lies, we now have a huge backlog that needs prioritization.

I know how to build my metrics, but I don't know what tools I should use to build them

Deciding a proper infrastructure is not an easy task, but as always, it should be driven by a business need. In this case, rather than the business demand for specific metrics, what we should be considering is our long-term, company-wide strategy, which should revolve around who will be working with data, how and when.

How much data do we need to store for analysis? How far will it grow? How fast? Do we envision a multiplicity of users accessing different layers of information to build over, a limited set of users enabling established, company-wide metrics, or both? What type of analysis do we want to perform? Will the decisions be taken close to the business facts or is a gap acceptable? How long is that gap? Is our most valuable data structured or unstructured? How much transformation do we need to apply to our raw data for a business user to find it useful?

With those questions answered, you'll have in your hands the information needed to decide what type of repositories, integration and analytics solutions you'll need. Whether it is a data lake, a data warehouse, a data streaming solution or a classic ETL, if there is awareness of what is needed, a good decision can be made.

Tool selection itself, will depend on several factors. Once you know what you need, it must be balanced with cost, the tools you already have and the interoperability with the ones you plan to acquire, and the existing IT know-how that your company has.
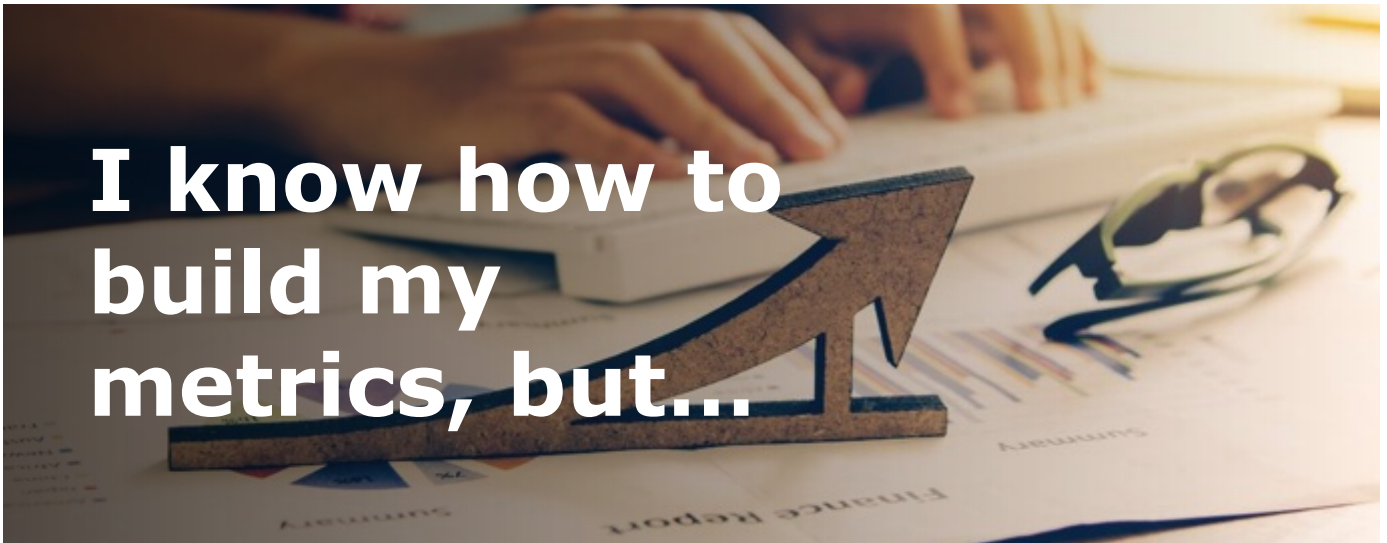
I know how to build my metrics, but...

...I don't have all the resources to calculate them. Some metrics we planned will be easy to calculate and meaningful to the company. We will call those, the Quick Wins and those should be the first we create, since they will provide us with success cases that turn into good faith from the company and will allow us to undertake other, more complex tasks.

Other metrics require large amounts of time to build but will be as important, if not more. We will call those our Major Projects and will follow our Quick Wins now that we've gained some credit with our first delivery.

Some metrics demand little time to build but won't add much value to what's critical to the organization right now. We will call those our Backlog, and they are the perfect task to do between projects.

Finally, some metrics will consume a lot of time and won't add much value to the organization. We will call those Time Sinks, and we won't work on them until priorities change their criticality or there is an easy way to deliver them. This might seem counter-intuitive, but some metrics might become easier to solve once there is a solid foundation. **Likewise, a change in the company's strategy could turn a secondary metric into a key one.**

06.

I built the metrics ... is that it?

If only… Building the right metric is one side of the coin. Information is as good as its users. Metrics need to be embedded into the company to unleash their true potential. A perfectly calculated, high quality metric with no users, is a poor metric. This is a point of struggle for a lot of organizations, sinking tons of money into information that just sits there, adorning dashboards no one ever reads.

This poses a great challenge to many organizations that can't seem to take the final step into a digital transformation. The causes for this are plenty, but in my experience, these have been the most common:

1.
Users prefer to do their own metrics in excel and present them to the CEO – This might sound weirdly specific, but I have seen many incarnations of this issue which covers at least two big problems: a. the CEO accepts parallel reports, which means there is no C-Level support for a unified, company-wide Data initiative, or b. the users are unsatisfied with the metrics given to them and decide to fix those issues themselves, which in some form or another is covered in our next issues

2.
Users don't trust Information because it was of poor quality at one point – Trust is something that takes a lifetime to build and only a second to destroy. However, if we accept that perfection is unattainable, we can at least have a solid structure that allows us to answer to any problem a user finds in a timely manner if the mistake is simply in the calculation. If the underlying issue is that the source data is unfit for use, refer to the quality rules in **STEP 3**, quantify it and make them available to users so they can know beforehand whether they can trust the data or not. If bad quality persists, you have the perfect justification to deploy a project that remediates data quality at its root cause.

3.
Users say the information is not useful – If a metric was not business driven, it had no reason to be published. It might sound harsh, but that's why **STEP 1** is critical. From its inception, the number must be viscerally linked to a business need, otherwise is just a complex algorithm that devalues the rest of the information. In these cases, less is more. A small dashboard with few but precise and valuable metrics is better than a hard to read one where the information was not needed in the first place.

4.
Users say the information is hard to read – This is where visualization comes into play. We need to walk our users that extra step, explaining what an upwards line means in different contexts, but we must make sure that we've chosen the right visualizations as well. Our first attempts should be kept simple enough that our users have to take a step instead of a leap. If they prefer raw numbers instead of a glimpse into a graphic, we might want to get feedback from them.

I went through all those steps, what now?

If you defined a business need as a driver, translated it into a metric, catalogued the sources and terms needed to build it as well as the rules to endorse its quality, then prioritized it and delivered it to happy business users that now work with those metrics in their everyday tasks, you did a GREAT Job. And likely you are not wondering where to start, but rather where to go next.

The two main courses of action next are to grow and experiment
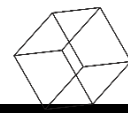
Growth is best experienced where the foundations are solid. Most steps are not a single area effort, but rather a dialogue between business and IT, co-creating value out of the data. Make sure those spaces are frequent, with clear objectives and active participation. Lay the structure so business becomes accountable and an active creator of value through data. Steps 1 to 5 are not a one-time thing. There are probably a lot of lessons learned and initiatives that follow your first quick win, so everything can be subject to re-thinking if there is a set place and time to do it.

Experiment with the newest technologies. If your data is properly catalogued, available and the drivers are clear, there is room to think outside of the box in terms of metrics and try a more cutting-edge approach such as AI to improve a business process or make better predictions. Some drivers you might have thought of at STEP ONE were wishful thinking 10 years ago and now are the order of the day. Are you willing to miss that train?

baufest
evolving business together

baufest

evolving business together