5 TIPS FOR EMBEDDING METADATA INTO IT OPERATIONS



If you're struggling to come to terms with implementing data initiatives in your company, then establishing processes for metadata (data *about* data) might feel a bit … overwhelming, to say the least. True, metadata may not be the sexiest part of a data initiative, but it offers the same benefits of data: providing insights, improving efficiencies, and testing new ideas. Leaving metadata out of your overall data mission is the same as leaving extra cash on the table.

If the real message of 2018 is that data shouldn't be left just to the IT folks, then data should be used for and by everyone. The best way to support this is by growing a company culture that cares about high quality, trustworthy data – and part of that is using metadata to maximize the potential and value of your data.

What is metadata?

The term *meta* is anything that reflects or refers back to itself. So, metadata is literally data about data. If you're not a data enthusiast, then metadata sounds even less exciting. But here's the thing: it's certainly technical information that doesn't seem necessary at first glance, but metadata *is essential* to data, especially if your company relies on data.

Metadata can describe a range of things, from the information embedding in a single file to information about an entire dataset. For instance, in a single digital photo, embedded metadata (that's often automated!) may include time and location of photo, GPS coordinates, keywords, original file name and file type, and the photographer/author, and even the person who owns the

rights to it.

Metadata about datasets and databases can include even more technical pieces, like how the data is digitally structured and formatted, where it's physically stored, who can access it, and what table and field names are used. Metadata can also incorporate business details such as a meaningful context of what the data is and which teams create or own it. Operational metadata also ensures data quality and data lineage, which can help identify and correct errors and enable automation over the lifecycle of data: its creation, its transformation, its movement, etc.

When metadata is embedded and integrated into several areas of the business, more people benefit from the data itself.

Benefits (and difficulties) of embedding metadata

The benefits of metadata are much like the benefits of data itself:

- Embedded metadata helps ensure consistency across your data.
- Metadata classifies information in a way that's easily sortable and findability. (Ideal use of metadata classifications can make it as easy as an internal search a la Google to find just that image you need in a storehouse of thousands.)
- When users can find data efficiently and correctly, the overall trustworthiness of your data increases.

These benefits support the ultimate goal: when metadata is managed across business, technology, and financial sectors, you no longer need to reinvent the wheel with each and every new data initiative. Metadata clarifies what data exists across the entire company, so more people can access the results of data analytics and put the data to work in ways it hasn't yet been tested.

And therein is the biggest resistance to metadata. When data isn't in harmony across the entire company, your data efforts are not being maximized. Harmonious data is the opposite of each team or department collecting, storing, and analyzing their own data. Harmonious data is the opposite of a team using its own descriptions and terms to describe what's being captured.

When data isn't in harmony, there's a company-wide lack of knowledge around what data exists. This can result in extra unnecessary work – gathering new data for a project when applicable data already exists. Or, people aren't sure whether they can trust the data if it's the lingo of another department.

5 tips for embedding meta data

Embedding metadata into IT operations doesn't have to be hard – and these five tips make it easier!

1. **Be consistent.** While consistency is usually an overall best practice or a good intention, it's actually essential to metadata. One major problem with company-wide data is that different stakeholders, departments, and even management talk about the same people, processes, and data in different ways. For instance, customers might be called consumers or clients or they may have more specific names altogether depending on who's doing the talking. Before you even begin embedding metadata, make sure you have a company-wide way to understand terms, meanings, and calculations. (Individual departments can continue using their own lingo, but for purposes of tracking and understanding holistic data sets, consistent

terminology is key.)

2. **Start small and grow big.** Trying to capture, integrate, and put to work metadata all of a sudden is a surefire route to failure, especially if you're talking about truly big data. With your next data initiative, add in a metadata process for it: start with small, easily trackable details, and you can learn as you go. As projects continue, your understanding of how best to incorporate metadata will grow.
3. **Clearly define processes for each step of the lifecycle.** Some data experts say start with the end in mind – what users will need to access this data, what initiatives require what data, etc. This is useful but it isn't holistic enough. Instead, consider the processes that data undergoes at each step of the lifecycle – acquisition, modification, retention, and deletion. This will help enact specific, efficient processes at each stage.
4. **Embed responsibilities into existing roles instead of creating new roles.** Just as you're starting small with your metadata initiative, start small with data responsibilities. There's no need – yet – to create a metadata steward or a metadata scientist. (This can also have the opposite effect of saying only certain people are responsible for metadata.) Instead, embed these responsibilities into existing roles, so that more people feel responsible towards data and metadata.
5. **Automate your metadata.** This one nearly goes without saying. Just as you're looking to automate your data collection, cleansing, transformation, and more, automating your metadata offers the same benefits: reduction of human error, improved processes, and increased trustworthiness.

Lastly, if you decide that a vendor solution is a better option for getting started with embedding metadata, remember that you can use the vendor's software without fully adopting the vendor's data theories. After all, you know your business best, so focus on the metadata you need.