

Justifying an MDM project to the business

Justifying an MDM project to the business

Let's start by approaching this from a different angle and asking "How can we justify not having access to high quality data in our business?" With this in mind, it's fair to say that high quality data is a critical element in operating a business. The truth is, we don't have unlimited time, talent and money, and you simply can't solve everything at once. You need to break these large projects down into consumable chunks. Every business manager should think "We simply can't wait to fix our data quality, and now is the right time to justify high quality data in our business".

There's now a high focus on data science, machine learning and artificial intelligence to help businesses guide data driven insights. However, this has highlighted that most businesses don't have the right data foundation to properly activate these new services.

At CluedIn, our mission is to help businesses simplify the challenge of taking the data throughout their business and raising the quality of it to make it "ready for business insight". We recently worked with one customer who had been confidently reporting to the business stakeholders that they sold to 4,632 cities around the world. We helped build the right data foundation for them, and the number of reported cities dropped to 1,591. The data was the same, but now it told the right story to the business.

So, let's tackle the same question from a slightly different angle - "How "do we justify investing in Master Data Management (MDM), Data Quality (DQ) and Data Governance (DG)?

The clients we work with have mandated that good data quality is their number one driver in improving business value back to the company. They come to us because they struggle with their past MDM, DQ and DG implementations. If you ask the main industry analysts, they all agree that MDM, DQ and DG are harder to establish than a Data Lake or Data Warehouse.

So why is this?

A traditional MDM implementation is generally considered successful if you can get a single domain operational within 6 months. It takes so long because there is a substantial emphasis on human interaction. Breaking down these human components, you need to model the data, identify, and develop the right business rules, and then build a complex ETL (Extract, Transform and Load) pipeline to make it all work. It's like trying to fit the "conventional square peg into a round hole".

A traditional MDM simply doesn't work in today's modern environment.

CluedIn's focus is to provide a simple way to justify the costs involved in a modern MDM solution by implementing an innovative "zero modelling" approach. You don't need to build your data model upfront, and hope your data fits in. We help accelerate your data ingestion so that you can start gaining data driven insights straight away.

Our approach allows a business to onboard new data environments through a scalable integration pattern. We know that you will encounter some tricky modelling situations where a traditional MDM solution simply doesn't work, and we cater for this. Our scalable on demand approach allows our customers to enable growth from 5 data sources to 25 data sources or even 250 data sources that were never meant to work together. The best part about this is that CluedIn can automate a significant part of the process that is easy to maintain and adapt as new sources come in.

We've learned from the project management industry that it is far more efficient to focus on small, gradual wins, instead of a grandiose "one and done" approach. In fact, Gartner states that over 50% of all MDM projects fail, as the business hopes everything comes together.

Data is a living, breathing, and forever morphing business asset.

Focusing on small, quick MDM wins that solve individual business challenges provides a more reliable Return on Investment (ROI). CluedIn helps break down a large MDM project into smaller business use cases that not only accelerate the implementation, but also the business ROI.

One of CluedIn's guiding principles is that "it's okay to let data out that has problems". It takes time to integrate, clean, train, enrich and de-duplicate data, and sometimes the business simply wants access to the raw version - like Twitter and Facebook feeds. Where possible, we provide tools that can identify common issues, like spelling errors, and automatically fix them as they come in to CluedIn.

Our approach of getting business use cases solved, operational, in production, and adding value sooner is more efficient than simply waiting for all the data quality issues to be solved before the business can use it. By using this approach, we can identify common data problem patterns more effectively. We understand this doesn't work for all parts of the business, especially when it comes to Personal Identifiable Information (PII) like credit card information. By understanding these constraints, we can easily apply the right business rules at the right time for each data point.

The more traditional approach means you have to go to every business unit and ask them what they "consider to be bad data quality" and then apply complex business rules that could tally up in the thousands. In our experience, we know that while there are similar data quality issues with most companies, there are still plenty of data quality issues that won't fit easily into a business rule. This further delays a traditional MDM project and increases its risk of failure with the heavy manual human interventions.

The time taken to generate all possible data quality issues upfront is nearly impossible to achieve, and often will delay a project by several months.

At CluedIn, we take a different approach that we believe is smarter and more scalable.

As you bring in data into CluedIn, we use technology to do the manual, repeatable work for you. This includes tools that focus on anomaly detection and smart text processing. With this approach, a data steward can simply accept or decline it. By using this approach, we believe the technology can pick up 80% of the challenges and help discover the underlying issues with the final 20%.

Our mantra is "it is better to make a process obsolete than to make a process simpler" and by doing so, we remove the underlying data problem in the first place.

Our pricing model is built on a simple premise - based on how quickly you want to clean, train, enrich and de-duplicate your data. We focus on the number of computer processing cores you need to get the job done, instead of worrying about the number of data points, attributes, users, or sources. This allows CluedIn to easily scale with the business and is more cost effective than a traditional MDM solution.

