

5 DATA MANAGEMENT STEPS YOU NEED TO TAKE NOW

Thinking of your organisation's most recent data and analytics projects, what percentage of the total time was spent finding and preparing the data compared to the time spent extracting value through visualisation, reporting, or analysis?

WHAT CAN MEDIA COMPANIES DO TO REDUCE THE TIME SPENT ON FINDING AND PREPARING DATA FOR ANALYTICS AND AI?

A common axiom in data science is that in most cases, data scientists report spending around **45%** of their time just on data preparation tasks. And **63%** of employees report they can not gather insights in their required time frame.

1

Evaluate putting your data in the cloud

2

Consider a layered approach

3

Use a meta-data driven method

4

Pursue automation

5

Contemplate compliance



EVALUATE PUTTING YOUR DATA IN THE CLOUD

Businesses need to make today's decisions with an eye to the future. For too long, businesses have rolled out IT initiatives only to see their new advancement become outdated on day one. The idea of future proofing, along with the uncertainty of predicting needs for data volume and technology progression, is pushing companies to the cloud. Moving to the cloud also helps companies align their modern toolset with their approach.

TIP:

Start out small. Develop a lake in the cloud, add a few tools, test, and experiment. Make sure what you design can scale to production. Azure is great for helping you build out a small environment suitable for production – in the cloud.

CONSIDER A LAYERED APPROACH

Different types of users require that data be made available for analytics and AI in different modes. For instance, data scientists typically require raw data that has not been cleansed, but data analysts want to work with data that has been consolidated, rationalised and cleansed. Business users of visualisation tools need governed semantic models with documented data sets.

TIP:

A layered approach means creating a data lake from a variety of data sources and then building one or more data warehouses from the data lake. Semantic models can be produced from either the data lake or data warehouses for self-service analytics. By using this layered approach, you ensure that everyone working with analytics or AI is using the same data.



USE A META-DATA DRIVEN METHOD

In a report by Gartner – “Predicts 2019: Data Management Solutions” – the authors encourage readers to make the strategic planning assumption that “By 2022, organisations utilising active metadata to dynamically connect, optimise and automate data integration processes will reduce time to data delivery by 30%.” Gartner says “Most organisations continue to struggle with provisioning integrated and curated data from heterogeneous data sources to data consumers, to support their data and analytics use cases.”

TIP:

Writing code to extract data from operational systems is time consuming, error prone, and provides no documentation other than the code. Meta-data driven methods will auto-generate SQL code for structured data and document where the data came from. If the data was cleansed, merged or altered in any way – that is documented too..



PURSUE AUTOMATION

Companies are recognising the benefits of data automation and moving towards automating everything that can be automated within the data life cycle. There's a disconnect between available resources and the number of projects on IT's plate. Add to that the fact that projects, especially data projects, are becoming more complex, and it's clear that companies are faced with a predicament. Automation helps solve the constraint of a lack of resources.

TIP:

It's time to stop talking and start doing. Organisations don't need to lay out a full business case and ROI calculation for using analytics and building out their supporting infrastructure before getting started. You won't always know in advance what you might find that could be useful, and you certainly don't know what data quality issues you might have. Make the commitment and move towards automating complex, time-consuming, redundant tasks.



CONTEMPLATE COMPLIANCE

If you are not afraid (of compliance issues for your data) – you will be. GDPR has been in effect in the European Union since May 2018 and as of the end of 2018 many believe that 50% or more of companies based in the EU are not yet compliant. Privacy and security rules like HIPAA or PCI may impact your data depending on your industry and many expect that GDPR-like rules will be implemented around the world.

TIP:

Contemplating compliance today means being able to document data lineage and access rules in metadata about the data you have placed in data lakes and warehouses for use in analytics and AI. The layered approach mentioned previously makes this easier..



Want to know more? Let's talk, or even better show you.

Visit www.touchstonebi.co.uk

Email info@touchstonebi.co.uk

Call 0207 121 4701