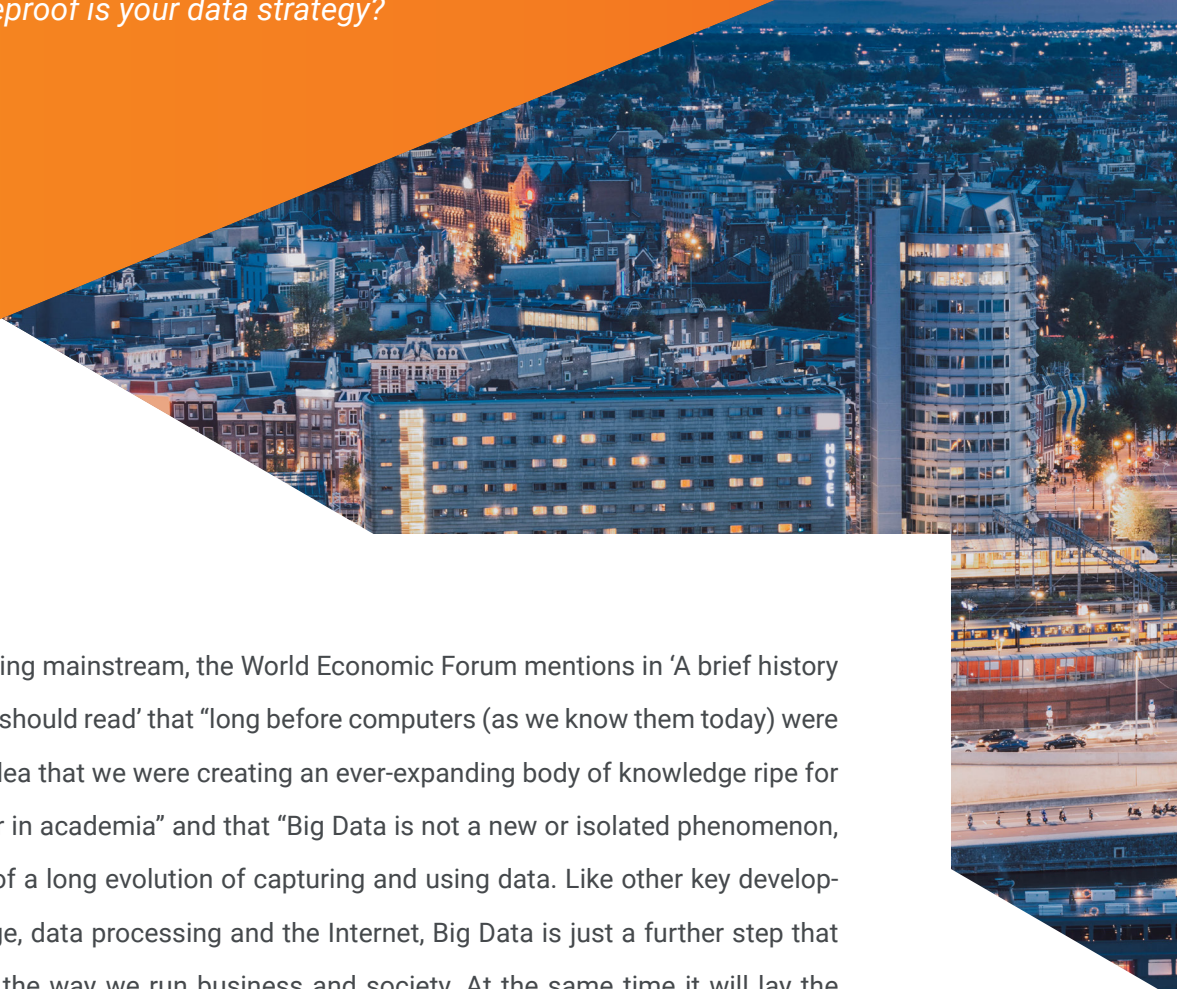# Checklist

For Building a Modern Data Estate

An important guide for any organization who wants to build a modern data estate but does not want to spend too much time or money doing so!

TIME**X**TENDER

Maturing towards a modern data estate in 8 steps

Data, data, data. Data everywhere. Data from enterprise systems. External data. Customer data. People data. Data from social media channels. Raw data. Unstructured data.Semi-structured data. Is your organization able to unlock value from all these data sources? Can everyone in your organization access the realtime data they need to make data-driven decisions at the right time? How mature and futureproof is your data strategy?

A brief history

With big data becoming mainstream, the World Economic Forum mentions in 'A brief history of big data everyone should read' that "long before computers (as we know them today) were commonplace, the idea that we were creating an ever-expanding body of knowledge ripe for analysis was popular in academia" and that "Big Data is not a new or isolated phenomenon, but one that is part of a long evolution of capturing and using data. Like other key developments in data storage, data processing and the Internet, Big Data is just a further step that will bring change to the way we run business and society. At the same time it will lay the foundations on which many evolutions will be built." This tells us that the abilities of an organization towards capturing data, data storage, data analysis and searching, sharing, transferring, visualizing, querying, updating of data, as well as compliance and data privacy are no longer a want-to-have.

These capabilities have become a foundation for success, today and in the future.

TIMEXTENDER

From a data warehouse to a data estate

"When setting up a data estate, think of it as building a house. If the foundations are shaky, or if you get to the first floor and realize you forgot to install any pipes, then you're in trouble. Solid foundations are absolutely vital to a data management platform, so make sure you establish a firm basis of support within the organization. This may mean that the implementation takes a little longer, but it will ensure you are truly prepared for the future by the end of the journey."

Most companies have already invested in their data environment by deploying a traditional data warehouse (DW). A data warehouse is a central repository of integrated data from disparate data sources used for reporting and data analysis. The traditional data warehouse however has limitations in functionality. Most data warehouses are typically updated as an end-of-day batch job, rather than being filled by realtime transactional data. And in a structured data warehouse environment work must be done within the framework of the created structure: static data sets with minimal ability to drill down.

Therefore more and more organizations are investigating the possibilities of developing a modern data estate. Or, as Forbes states in the article ' Why The Modern-Day Corporation Should Consider A Data Estate': "Businesses are scrambling to build a data infrastructure suitable for supporting their desire to mine the riches inherent in their data.

And they're doing so with data estates." The same article offers us a great definition of data estates: "A data estate is simply the infrastructure to help companies systematically manage all of their corporate data. A data estate can be developed on-premises, in the cloud or a combination of both (hybrid). From here, organizations can store, manage and leverage their analytics data, business applications, social data, customer relationship systems, functional business and departmental data, internet of things (IoT) and more."

Do you have a traditional data warehouse (DW) and have been wondering how to strategically expand its functionality and upgrade its performance? In this paper we discuss 8 steps to take when expanding and upgrading your data warehouse into a modern data estate.

01 Define your goal in terms of data and analytics maturity

How mature is your organization in the area of data consumption? What do you want to achieve with your data estate in terms of analytics maturity? In the article 'Take your Analytics Maturity to the Next Level', Gartner cites that

"In a recent Gartner survey, 87.5% of respondents had low data and analytics maturity, falling into 'basic' or 'opportunistic' categories.

Organizations at the basic level have business intelligence (BI) capabilities that are largely spreadsheet-based analyses and personal data extracts.

Those in the opportunistic category have individual business units that pursue their own data and analytics initiatives as stand-alone projects, but there is no common structure across them."

In this step you ask yourself what is currently working and what is not working. You describe the bottlenecks, pain points and the data silos in your organization.

At the end of this step you have a clear understanding of what you want to achieve, which pain points you are going to solve and where the low hanging fruit can be found.

02 Define the business needs of today and the future

"In a competitive environment, where data can make or break a businesses' competitive advantage, corporate success might very well be measured by the maturity of its enterprise data program" mentions the earlier cited article from Forbes.

In this step, it's not about IT. It's not about limitations, restrictions or data silos. It's all about a laser focus on business requirements. When transforming your data warehouse into a modern data estate, the biggest mistake one can make is a replication of the existing environment into a new environment. You need to ask the right questions to every stakeholder involved, from sales to marketing and from HR to operations.

What do they want to measure? Where will their business be headed to in the future? What trends do they see? How will digital transformation impact data insights? An HR-leader wants to measure traditional KPI's today, such as employee engagement, diversity or belonging. Trends in HR, such as the war on talent will however mean that HR needs insights into the individual skills and competencies of employees so that scarce skills can be quickly allocated to the most strategic projects.

Following your tour of the business it's time to prioritize these business needs, as you will want to start out small.

03

Describe the core business and data processes

You've defined your goal in terms of data and analytics maturity. You've prioritized the business needs of all your stakeholders. Now it's time to identify the data sources you have available – and which data you want to initially find a home in your data estate to respond to the highest priorities from your list from step 2. In this step, it is both about business processes and data processes, since they are connected. You look at a certain data point, such as customer data, and then you define the relational data models. How is this customer data is used in your business processes? The same applies to transactions, products and more.

04

How will the data be accessed and by whom?

Employees are different in the data they have access to (security) and in the way they access the data (tooling). In this step you describe your security strategy and the tools for analyzing, reporting and visualizing data. The first question to answer: How are you going to connect to the data sources? What connectors do you need? How often can you read data? What kind of data do you get? What metadata is available? How often is data updated and how often do you have access to data? The second question to answer: how will you manage security and access rights?

How people access the data

Consider the data consumers you want to serve, and how you want to serve them. You want to provide self-service BI for different types of consumers, ranging from power users such as data scientists, data miners, AI and ML algorithms to business users working ad-hoc with data and creating new reports, and casual users waiting for the routine reports and updated dashboards.

The data people can access

In this step you define roles and groups so that, building your data estate, you can identify users and access rights. This ensures that authenticated users only access the data, tables or columns they are authorized to see.

TIMEXTENDER

05 Define your architecture

Data needs to be extracted, processed and refined to be useful. And just as oil can be refined into different types of fuel, data can be prepared for different uses when it comes to analytics and artificial intelligence. In this step you describe how your organization chooses to prepare data for these different uses, from reporting to analytics and artificial intelligence. Most data estates are split in 3 distinct layers: the data lake, the data warehouse and the data marts.

The final result is an integrated architecture that significantly reduces costs, accelerates time-to-value and supports your data compliance needs.

Data lake

This layer is primarily for power users such as data scientists, who perform various types of analysis on raw data to look for anomalies and patterns, and eventually perform machine learning. This layer enables quick ingestion of raw data from all data sources and into Azure Data Lake or a SQL Database.

Data warehouse

Raw data isn't the best choice for business users, such as business analysts. These users need data that has been cleansed, enriched and rationalized – in a modern data warehouse. In a layered data architecture, this data warehouse would be sourced from the data lake – but placed in a SQL-based database with semi-structured data transformed into structured data for analysis.

Data mart

The data mart supports common users by delivering relevant datasets from the data warehouse, enabling self-service analytics across multiple analytics tools for line of business or function specific views, so that business users can explore data safely and efficiently.

06 Cloud, On Premise or Hybrid

Where in real-life the foundation is essential for the building that's constructed, in data-life the foundation of your data estate is equally important. You don't want your data estate to end up like the leaning tower of Pisa, where your data estate becomes a costly affair to maintain. Consider the pros and cons of cloud, on premise and hybrid. There are great cloud solutions available on the market – such as Microsoft – but "Cloud should be thought of as a means to an end. The end must be specified first", says David Smith, Distinguished VP Analyst and Gartner Fellow Emeritus in the article 'The Top 10 Cloud Myths'.

07 Selecting your construction partners

In this step you select your construction partners. Which software will you use for your data estate? Who will you build the estate? And how will you maintain the estate?

Data management and automation software

You should select the right software platform for today and the future. You will want to ensure that your data estate is built in with an integrated data management platform that is completely independent from developers, data sources, data platforms (SQL Server, Azure SQL, Data Lake, Synapse), front-end tooling (Power BI, Qlik) and deployment model (on premise, cloud, hybrid). You should be able to expedite development with automated code-generation, freeing data engineers to focus on data quality and business requirements and limit the required number and types of high-ly skilled resources by using a single tool to build your data lake, data warehouse, and data marts. Last, but not least, you will want to ensure you data estate is 'future-proof' meaning it is fully scalable and ready to adopt future releases without rebuilding.

Deployment and maintenance partner

Will you deploy and maintain the data estate yourself? Will you consider a deployment partner and then take on the maintenance yourself?

The latter is what many organizations opt for – as they want to ultimately be able to take the control of their data in their own hands, and not having to depend on a business partner. However, whatever you decide – consider a partner with experience and a partner you can trust, as they will be deploying a future-proof foundation for your most valuable asset: data.

08 Think big, start small and act agile

Have you followed all these seven steps? Then it is very likely that your data estate will help drive innovation and that you will be deploying a scalable, future-proof data environment. It is key to start small. For example, by developing an estate in the cloud (which you can scale to production), with only a couple of data sources, a few tools, and then testing and experimenting. Step by step you'll then help your organization to transform into a data-driven organization.

TIMEXTENDER

THE AUTOMATED WAY TO BUILD A MODERN DATA ESTATE FOR ANALYTICS AND AI

TimeXtender is an integrated data management platform that helps implement and operate data lakes, data warehouses, and data marts – without writing code – automating the process of getting documented data in the right place, in the right form.

With TimeXtender you can transform your DW to easily and completely integrate into a modern, automated data-estate platform. This means that initial costs to build a modern data estate - for an organization with an existing DW - will be very economical and the time to transition will be high-speed. And when transitioning from a DW to a modern data estate with TimeXtender, the platform helps you with each step on your journey:

Orchestrate building and maintaining the complete data estate – not just a narrow portion – with highly integrated capabilities.

Manage development, test and production environments without rewriting code to switch data platforms – run dev and test in less expensive environments with different latency requirements, and ready for analysis.

Define data marts once and generate multiple models for self-service analytics (metadata-based modeling)

Easily add additional data sources and elements and avoid building custom datasets for each data request – speeding maintenance and reducing IT backlog

Automate the generation of complete documentation from project metadata

Manage development, test and production environments without rewriting code to switch data platforms – run dev and test in less expensive environments with different latency requirements.

Easily change deployment models (on-premises, cloud, hybrid) or data platforms (SQL Server, Azure SQL, Data Lake, Synapse) without rewriting code – ensuring your data estate is fully scalable and ready to adopt future releases and new technologies.

Identify and secure sensitive data while maintaining and documenting access rights – supports compliance with internal and regulatory requirements.

Work with your trusted partner – as TimeXtender avails over one of the broadest partner networks in the industry.

TIMEXTENDER

www.timextender.com - info@timextender.com