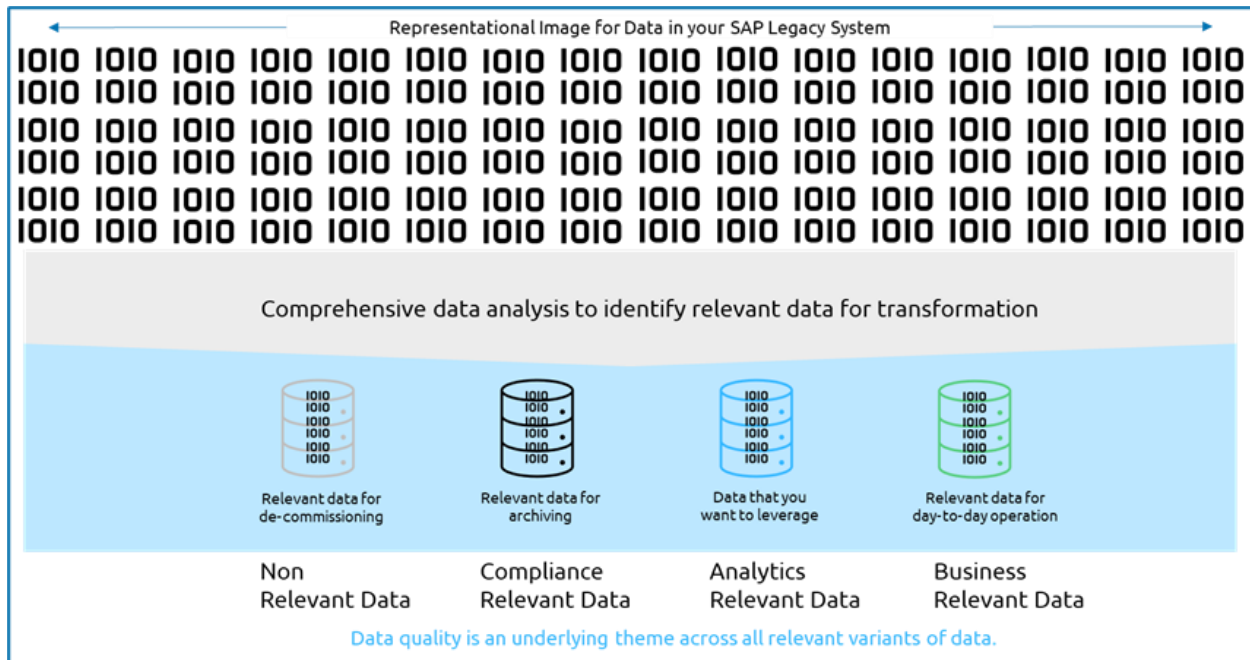# Managing Data in a Large-scale SAP Transformation

Managing data in a large-scale SAP transformation can be challenging. It is vital to comprehend the data in the context of a big-scale transformation and act accordingly. When you decide to move either historical data (org unit wise or status-specific) or open documents with master data from your existing SAP ERP system to the SAP S/4HANA target system, you must ask yourself what happens to the data that isn't moved. How can you still use it in your transformation strategy? Since 'data is the new oil', we simply cannot afford to disregard the data that is left behind. For any large-scale transformation project, it is crucial that business operations continue uninterrupted during the implementation of essential design concepts such as lean core, cost-effective target infrastructure, and targeted data migration.

Our observation based on numerous large-scale transformation project experiences is that the emphasis is placed on developing future procedures and selecting important data for day-to-day business operations, but not on the value that can be derived from data that is left behind. Typically, such data is either archived for compliance and reference or de-commissioned. This is a suboptimal strategy that necessitates careful planning well in advance to exploit all your data, regardless of whether your transformation strategy involves new implementation or selective data transition.

As part of transformation, the data can be evaluated to classify them into distinct buckets. The data that will be migrated to SAP S/4HANA in the future is business-relevant data. The remainder of the data can be divided into data relevant for analytics, data relevant for archiving, and data relevant for decommissioning.

Representational Image for Data in your SAP Legacy System

Comprehensive data analysis to identify relevant data for transformation

Relevant data for de-commissioning	Relevant data for archiving	Data that you want to leverage	Relevant data for day-to-day operation
Non Relevant Data	Compliance Relevant Data	Analytics Relevant Data	Business Relevant Data

Data quality is an underlying theme across all relevant variants of data.

It is crucial from a business case and operational standpoint that only necessary data is transferred to the new SAP S/4HANA system. Since the specifics of this choice will vary from one client to the next, the criterion for determining what information is useful for analytics will also vary. You don't want to relocate the last 10–20 years of data in the target system as doing so would be time-consuming and expensive. One client may require an equipment dashboard that combines data from the current system, along with the remaining 10 years of data. Other clients may be required to submit the maintenance history of components or product batch details to maintain legal or safety compliance. There is a need to evaluate a broader scope of analytics-relevant data, such as process-specific data over the past x years, for reporting or for feeding it to intelligent applications for improved forecasts and judgments. We have seen these requirements coming up frequently from customers in aviation, pharmaceutical, retail and manufacturing industry.

As part of the SAP ERP transformation journey, there is a requirement for a comprehensive approach and a cost-effective and efficient solution to address the complete data scope. This approach should enable clients in identifying the data that is relevant for the target infrastructure and conducting a deep dive into the processes to uncover analytics-related data. Clients must be able to choose analytics-relevant data based on their requirements, therefore the solution must be adaptable. Similarly, to business-relevant data, this data must also undergo transformation. Other significant considerations include limiting infrastructure expenses to manage analytics-relevant data, the data format for storage and access via an analytics tool, and the time-bound transfer of analytics-relevant data from the operational SAP S/4HANA system.

As part of our sustainable enterprise transformation strategy, Capgemini has established an analysis framework for advising clients on data management in large-scale transformations, as well as a solution to harness analytics-relevant data for compliant and efficient business operations. Our solution is built leveraging tools, products and solutions from our partners and is based on the following guiding principles:

1. To provide recommendations on data

2. To allow flexibility in data selection

3. To transform the data as necessary for both business and analytics relevant data

4. To store the data in a cost-effective format in the data lake

5. To provide the transparent access to the data (from either SAP S/4HANA or Data Lake) to the end user

The data transformation for both business and analytics relevant data is performed using cbs Enterprise Transformer Software®. cbs is our global and strategic partner for Selective Data Transition led SAP S/4HANA transformations. The data is stored in Microsoft Azure Data Lake for transparent access by end users in analytics report or within an application.
Whether you are building your roadmap for transformation or have already begun or completed your journey, it is never too late to analyze the data that is left behind to leverage the same for your business.

If you want to learn more about our data management experiences from 500+ SAP S/4HANA transformation projects and establish a compliant data management strategy and solution in your context, please contact us. Please contact Mr. Prabhakar Lal [Email: prabhakar.lal@capgemini.com Phone: +49 151 27729050] if you have any questions.

You can find this blog and any future updates on the mentioned solution here.