



# SKELETONS IN THE DATA CLOSET

Key considerations for archiving legacy data

White paper

# CONTENT

Preface ..... 3

Introduction ..... 4

The exponential growth of data and changes in the enterprise ..... 5

Keep data, or forget about it? ..... 6

The price of data ..... 7

Example case: SAP consolidation in the consumer goods industry ..... 8

The wild west of systems and data ..... 9

Maintaining the status quo ..... 10

Preserving data – without the legacy system ..... 11/12

Clearing skeletons out of your data closet ..... 13

## PREFACE

At many companies, a legacy system is a place where data goes to rot and die.

Although most of these systems are practically never used, the data they contain is still extremely important. Compliance alone requires many types of data to be stored for long periods of time.

This white paper discusses the necessity of retaining data and the options available for decommissioning legacy systems without losing access to the data they contain.

This document was created by T-Systems with the assistance of KPMG. Both companies contributed their extensive experience in system and data archiving. Nevertheless, this paper only provides a synopsis of the topic. It is not an exhaustive resource, but rather a snapshot of the situation at the time of publication (December 2018) and cannot substitute for the legal counsel a specific case may require. Despite the great care taken in drafting this document, the publishers assume no liability for its content.

# INTRODUCTION

This document is all about data.

- **Why is data important?**
- **What are the costs and risks involved?**
- **What laws regulate the deletion and retention of data?**

This white paper offers several options for banishing the skeletons from your data closet. We cover measures to minimize the expense of maintaining legacy systems as well as approaches for decommissioning those systems – without losing access to their data.

Why do so many companies have gigabytes worth of legacy skeletons in their closets? Many causes can contribute to the situation. These are some of the most common reasons:

## 1. Migration to a new system

This is something every company has experienced. The old system gets replaced by a new one. But the new system's functionality and data don't correspond 1:1 with the old system. This is particularly common when migrating ERP systems from SAP R/3 to S/4 HANA.

## 2. Mergers and acquisitions

Legacy systems are kept operational to ensure access to historical data whenever necessary.

## 3. Shifting IT strategies

Many companies transition from conventional on-premise infrastructure to the cloud or software as a service via a greenfield project that only includes the most essential data. But the organization still needs access to its historical data.

## The question that always comes up

Everyone wants to know: what's the easiest way to get rid of these legacy systems? And, is there any reason I can't do that? A number of obstacles can get in the way. Regulations that require the retention of tax documents, for instance. Product liability also requires the retention of production data for many years. And sometimes, systems are simply forgotten. Or the cost of maintaining them is underestimated. Knowledge of the available options can be lacking as well.



**It keeps getting more difficult to deal with all the skeletons hiding in an enterprise's data closet as IT permeates every facet of the organization. Each year sees an increase in the quantity of data that is generated. This deluge of information results in higher storage expenditures. In this way, the relative cost of managing legacy systems and their data increases over time.**

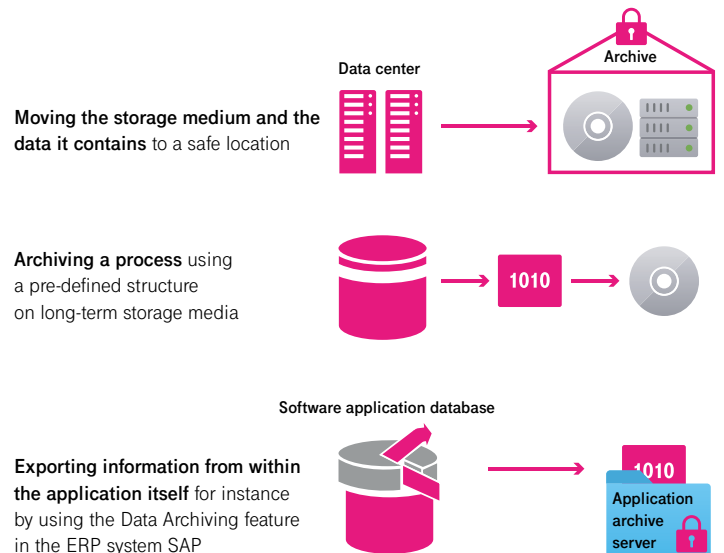
## DATA GROWS EXPONENTIALLY AND ENTERPRISES CHANGE

Data are indispensable when it comes to managing a business. Without data, a company would be flying blind. Over the course of the digital evolution, the quantity of data that is collected and stored has risen. Big data has become a part of many aspects of the enterprise including purchasing, sales, finance, production, and logistics. The enormous expansion of the volume of this data has resulted in a corresponding increase in the cost of the IT resources needed to store and retain it.

All the while, enterprises are experiencing accelerated change including more mergers, divestitures, and acquisitions. The new rights and responsibilities that these organizational changes bring with them make data an important topic. When one company purchases another, the acquired organization frequently has IT systems that are no longer needed. But tax regulations dictate that the data in those systems cannot be deleted.

When we consider this increasing volume of data being produced and the accelerating pace of organizational change, it's easy to see how large, complex IT estates form over time. Yet many of these systems – and the data they contain – are hardly ever used.

### The many meanings of the term “archiving”



## THE SIGNIFICANCE OF ARCHIVING

The term “archiving” is used to signify a number of different things. In general, people use the term archiving to talk about the long-term storage of objects or information. In a conventional city archive, for instance, climate control helps to preserve important documents in their original form. If a document is needed, however, it cannot be accessed immediately. An archivist must first determine the document’s physical location before they can retrieve it.

With digital archives, old city documents are preserved not in their original form, but as digital copies on storage media. So the system no longer revolves around the physical object itself, but rather the information it contains.

In the digital realm, the term “archive” has many different meanings. It’s used to talk about digital media such as tapes that hold data and are stored at a safe location. In other cases, archiving refers to a process for extracting structured data from one type of storage and transferring it to long-term storage media. And sometimes archiving is used to denote the extraction of information on the application level, such as the Data Archiving feature in the ERP system SAP.

In this document, we will use the terms archiving, digital archiving, and historization interchangeably to mean the following: Essentially, we’re talking about moving information from one place to another where it cannot be modified and can be stored for the long term. This does not presuppose any particular structure or format for the data.



## HOLD ON TO DATA, OR FORGET ABOUT IT?

Archiving enterprise data is necessary for many reasons. Commercial and tax regulations on data retention, product liability, pensions for federal employees, and technical reasons involving performance and storage are just some of the factors that can make archiving data a necessity. Regulations and standards in Germany, for instance, require information to be retained for up to 10 years in most cases. The applicable documents include the German Commercial Code (HGB); the EU GDPR; the tax code, and the FAIT 3 accounting principles.

In addition to these retention requirements, the “Principles for properly maintaining, keeping and storing books, records and documents in electronic form and for data access” (GoBD) also apply in Germany. The GoBD requires data to be complete and without omission so that all business transactions are represented. In addition, all steps during the creation and processing of business transactions must be completely traceable.

Archiving data in the form of extraction lists or archived business transactions represents a high IT compliance risk because complex business transactions are often very difficult or even impossible to reconstruct. These types of records may also contain gaps. Consequently, it appears appropriate to choose the most comprehensive and long-term archiving solution possible.

These considerations contrast with to the EU General Data Protection Regulation (GDPR). Effective as of May 25, 2018, the new regulation calls for “privacy by default” in digital archiving. One of the key stipulations in this legislation is that personal data must be deleted as soon as the right to store it expires. This includes creating plans for deleting data that is no longer subject to retention requirements and cataloging all stored personal information. Since archiving data often involves complex interdependencies and linked resources, it’s often difficult to store data according to expiration date and delete it once the deadlines have expired.

The potential fines for non-compliance are large and can be set as high as 4% of annual sales revenues. This makes it is absolutely essential for businesses to scrutinize the stipulations in the EU GDPR. Companies may wish to weigh the risk of fines against the expense of implementing privacy by default.

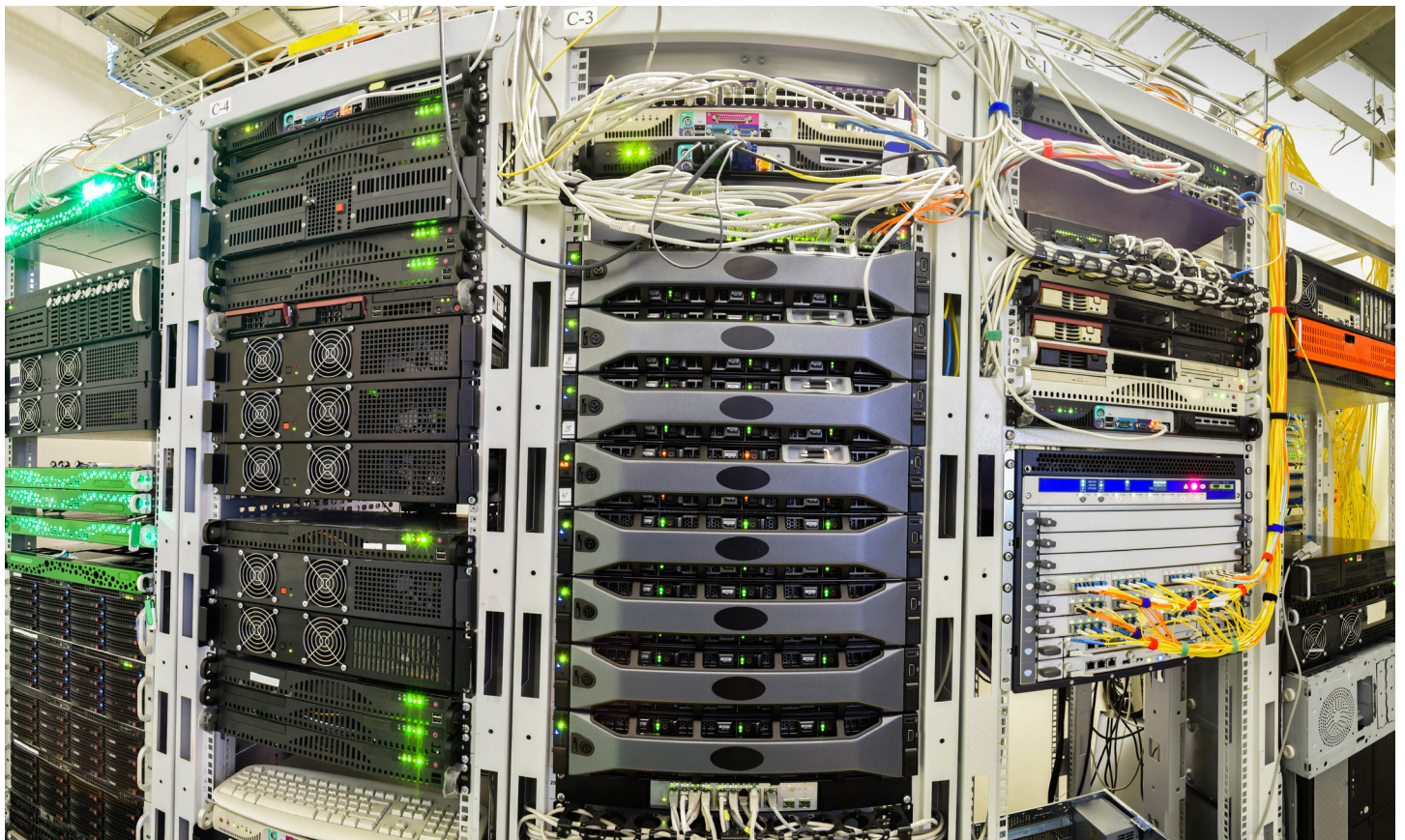


## THE PRICE OF DATA

Experience shows that operating legacy systems is very cost-intensive. When you consider the limited amount of time a legacy system is used each day versus the cost of operating it, you quickly recognize how expensive old data can be. Because all that matters is the data, not the deprecated features the legacy system offers. Nevertheless, legacy systems consume a sizable portion of operational IT expenditures (OPEX).

**An example:** A business decides to keep a legacy system around to provide access in the event an ad-hoc query arises. After a while, the company loses sight of how much it's costing them to keep the system running. Costly maintenance agreements with the manufacturer or software vendor are common in such cases. New systems with large feature sets often come with expensive maintenance contracts. The features may no longer be needed, but the contracts remain.

Replacement parts can also become very expensive in the event of a hardware failure – if parts are even available, that is. In some cases, experts with specialized knowledge must be kept on staff to operate the system. These costs are frequently ignored.





## SAP CONSOLIDATION IN THE CONSUMER GOODS INDUSTRY

A publicly listed German company that has been operating in the consumer goods sector for more than a century is facing a heterogeneous IT estate with over one hundred applications. The company manufactures and distributes products around the globe. A steady stream of international acquisitions is part of the corporate expansion strategy.

The IT infrastructure is currently in the process of being consolidated. At a cost of hundreds of millions of euros, the SAP and non-SAP systems of the various divisions and regional branches within the enterprise are being migrated to a single SAP platform. In 2014, the company began looking for cost-efficient and compliant options for archiving its legacy data.

After reviewing suitable products, the company decided to start with a Proof of Concept (PoC) and afterwards implemented a service offering from the T-Systems portfolio. The PoC phase was conducted using a complex SAP system. A number of international divisions were involved in the six-month PoC which included a comprehensive testing phase. KPMG actively assisted during the PoC phase, helping to assure the quality of the requirements specification, procedures, and execution.

Once the PoC had been successfully completed, the company began working with T-Systems to build the History Service and

completely historize one application. Then KPMG provided quality assurance to confirm that the historized data was complete and correct. This included verifying the data transformation, access to the legacy data, minimum requirements for archived documents, and retention requirements. Since over 30 applications were slated for retirement, the company formed a dedicated competence center in 2015 to establish holistic planning and controlling.

**By March of 2018, 10 IT applications had been decommissioned, validated, and migrated to a compliant historization service. This made it possible to shut down the existing ERP systems for good. By late 2018, an additional five to ten applications are slated for the historization service.**

The following aspects were essential for the company:

- **The very promising business case**
- **Significant reduction of risk, which the quality assurance services from KPMG contributed to**
- **The company no longer needed to maintain in-house staff with the skills to operate the legacy system**





## A JUNGLE OF SYSTEMS AND DATA

Never before have we seen such diversity and volume of data and its associated systems. Yet data and systems are key success factors for any modern business. The right data strategy for collecting, maintaining, and processing data can serve as an engine that drives core business processes. The way these processes are linked and interact with each other has a significant influence on an organization's agility.

One popular data strategy is known as bimodal IT. This approach involves splitting the IT organization into two separate tracks. The slower, more conventional IT and the IT fast lane. Conventional IT offers reliability and stability while the faster, agile mode of IT development focuses on innovation and differentiation. The objective of this strategy is to develop the capability to quickly adapt to changing demands without sacrificing the reliability of conventional IT practices.

The challenge enterprises face is finding the right systems (or system) to collect, process, and store all the different types of data they encounter. Key considerations here include cost, performance, and availability. And regulatory aspects must also be considered. The chosen solution also needs to fit the company's IT strategy.

Particularly for legacy systems, it's necessary to find a cost-efficient solution that fulfills both business and regulatory requirements. Since legacy systems from both conventional IT and agile IT approaches exist, the archival strategy must consider both of these modes of development. It can be necessary to consider different archival solutions for the various categories of data and systems.

**With a clearly defined data strategy including an archival strategy, you can ensure that all requirements are met for each category of data or system including costs, compliance, and business processes.**



## MAINTAINING THE STATUS QUO

One option for dealing with legacy systems is to keep them operating, but with limited resources. This allows users to continue accessing the systems as they always have. However, all the necessary operating processes must also be kept intact.

Normally, this option is only feasible if the budgets for ongoing operation of the system have already been planned and approved. This implicitly involves carrying the original budgets forward from one year to the next. It's easy to just keep doing this. At first glance, it seems like a low-risk option. The primary costs are clearly visible – so long as the software or hardware vendor does not demand higher maintenance fees for outdated software versions.

In addition, users are familiar with the way the old system operates and no workshops or user training are required.

**These advantages are however outweighed by a number of serious disadvantages.**

After a period of roughly 10 years, upgrades to the legacy system become unavoidable. This normally leads to a very high and unexpected cost hit. In addition, staff with the expertise to operate the legacy systems are required during the entire archival time frame. Employees often leave the organization during this period, however. This means their skills must be passed on to other individuals – at great expense.

Operating several legacy systems greatly limits the enterprise's flexibility when it comes to infrastructure. Outsourcing plans can become much more expensive, for instance, if a large number of older systems require maintenance. Because this situation limits the number of vendors who come into consideration. Special hardware, operating systems, and the applications themselves can make it impossible for sourcing providers to operate certain systems.

As these costs keep piling up, the company is also missing out on the opportunity to implement a cost-efficient archival solution and free up budgets for future IT projects.

**Our recommendation is to develop an archiving strategy for the organization that clearly defines how to handle archiving for all the existing types of data and systems. Having a guideline lets companies include the costs of archiving when planning to switch over to a new system or conduct a system integration. This makes it possible to estimate the total cost of switching systems including decommissioning the legacy system.**

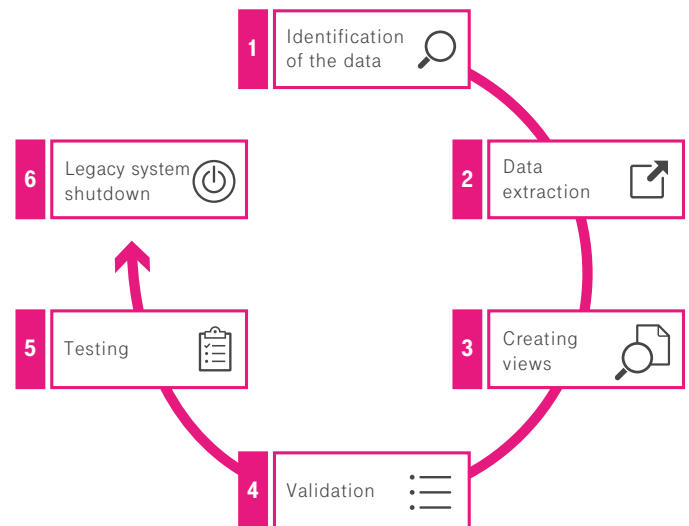


## PRESERVING DATA – WITHOUT A LEGACY SYSTEM

To historize the data and completely decommission the legacy system, the legacy data are migrated to a historization platform that is completely independent from the source system technology. Implementing this method involves the following steps:

1. Identification of the data to be archived
2. Importing the data into the historization platform
3. Generating views for the data in the historization platform
4. Validation and certification if required by an auditing firm
5. Testing whether the data are available in the event they are needed
6. Decommissioning the legacy system

The 6 steps of system archiving



The company can operate the historization platform itself or commission an IT vendor that offers a cloud-based solution, for instance. It's common for cloud providers to invoice based on actual use. Since access to legacy data is not usually required that often, the latter option can be very attractive. The more infrequent the access, the lower the cost. And there's no capital expenditure involved.

The historization platform can accept data from a range of different systems. Considerable synergies result when multiple legacy systems are scheduled for decommissioning.

A historization platform is specially designed to preserve data for decades. The application and database layer is normally decoupled from the hardware. The software components of the platform are based on open source products that are in widespread use.

### The key advantages of a historization platform include the following:

- The legacy system can be completely decommissioned and stops generating costs. The risk of data loss is eliminated. A backup strategy for legacy systems is often missing. Or, if backups do exist, they can't be restored in the event of a problem because the skills required to do so are no longer available within the organization. A historization platform provides up-to-date backup mechanisms capable of fulfilling data protection requirements.
- One of the most common risks of legacy systems – maintaining the necessary skills – is therefore eliminated. It's difficult to find – and hold on to – the right staff. With a historization platform, this situation is avoided by signing a service agreement.
- Expensive maintenance contracts and high operating costs due to outdated hardware are a thing of the past. Operating a historization platform is far less expensive in most cases.
- A historization platform can be used to retire many different types of systems.



Users receive a universal platform that contains all the historized data. Instead of operating countless legacy systems, you're left with just a single platform. Costs stop adding up. All that is left is the cost to operate the one shared historization platform.

Another advantage of integrating a large number of legacy systems into a single historization platform is the simplified, universal search interface. Whereas before, research required access to multiple applications and various user interfaces, now there is just one common historization platform. That's a huge advantage for users.

**In addition to these advantages, the following aspects should be considered when deploying a historization platform:**

- Investment is necessary, as it is with a virtualization solution (see previous section). The steps described above are part of the implementation project, and they generate costs. Normally, the company will see a return on that investment within just a few months.
- If the data have legal significance, the project should include (in addition to the data migration) certification of the project workflow, results, and the historization platform. These costs must also be taken into account when planning the project.
- Due to the fact that the application logic is not archived alongside the data, the data can no longer be displayed in the old system environment. To establish links between related information, for instance, the table structure of the database must be analyzed – which includes modeling the corresponding database queries.

**GoBD assessment**

This type of archive can result in GoBD compliance risks. A complete and fully intact data repository that allows all transactions to be reconstructed using the functionality of the original system is not available in this case. Since database archiving only encompasses the content of the database and not the application logic, reconstructing the original system functionality would either be very cost-intensive or impossible. The required ability to trace a transaction in both directions – from the individual receipt and ending with tax filings or the other way around – is affected. Complex database queries must be modeled to make this possible. The GoBD requirement for keeping receipts in their original format must also be taken into consideration for this method of archiving (GoBD, paragraph 131). An archiving strategy appropriate for the location the stored documents must be developed.

The relevant accounting documents and data can be stored directly in the database in the form of binary large objects (BLOBs), for instance. This allows them to be stored right alongside the rest of the data during the database archiving process, which means they remain accessible. If these accounting documents and data are stored in a different system or on a different server, however, they need to be backed up as part of a separate archiving process.

**EU GDPR assessment**

The EU GDPR requires personal information to be deleted once the right to store it expires (Article 17, EU GDPR). This functionality must be configured when customizing the historization solution (delete feature). Since these types of systems often have a complex data structure with many dependencies, the databases they are built on must also be analyzed and included in the deletion strategy. To comply with these EU GDPR requirements, an archiving plan must laid out in advance. Combined with the right technology, the resulting solution can then provide “privacy by default” as the legislation requires. Another thing to consider when developing an archiving strategy is the ability to put a hold on archived data to prevent deletion if data must be preserved to fulfill legal obligations (see Article 17, EU GDPR). Developing an authorization strategy requires limiting access to the data while simultaneously preventing unauthorized deletion from the archived database before the archival period expires.

**To sum up, although this method does require some investment, it offers a sound and sustainable foundation for keeping operating costs in check. Especially when several data sources are being migrated to the historization platform. The process of researching information is simplified and unified. The system can be operated as a cloud service.**





# CLEARING SKELETONS OUT OF YOUR DATA CLOSET

Cleaning up data stores is rarely one of the highest priorities in the enterprise. But being proactive here gives you two decisive advantages: significantly and sustainably lower costs and minimized risk. A typical approach can look something like the following:

- 1. Assessment:** First, obtain an overview of the current situation with respect to systems and data in the enterprise or business (initial assessment). It's important to establish or estimate what the current costs are as well as the expected costs if the current situation is maintained.
- 2. Requirements:** Collect all requirements that affect the storage and retention of data. The biggest influences here will likely be domestic tax law, commercial regulations, industry regulations, accounting standards, and laws that govern document and data retention such as the EU GDPR and Germany's GoBD.
- 3. Evaluation:** Analysis of the available solutions and assessment of the requirements that they cover (tax, legal, etc.), costs, flexibility, and additional internal requirements (such as IT strategy).
- 4. Definition:** Formulation of a global, enterprise-wide historization strategy including the necessary processes (checklists, etc.). The historization strategy can also consist of a combination of different solutions.
- 5. Communication:** Communicate the enterprise-wide guideline to the responsible employees.
- 6. Implementation:** Implement the historization strategy.

Thanks to years of experience applying this approach, T-Systems can gladly assist you.

The future will bring additional system consolidation and make many existing ERP systems obsolete. One example of this trend is the unavoidable upgrade path to S/4 HANA. This will cause a great many more enterprise systems to become superfluous.

The time to act is now, because without a clear strategy, additional costs and difficulties are certain to result.

A managed historization process reduces costs and frees up financial resources that IT can invest in innovative projects.



## Author

Harald Sondermann  
Senior Sales Manager

## Contact

## Published by

T-Systems International GmbH  
Hahnstraße 43d  
60528 Frankfurt am Main  
Germany