

TOLERANT Whitepaper

Mastering data quality

Challenges in ensuring data quality

Chapter 2.2 from the book entitled "Mastering data quality: How TOLERANT Match makes the difference" by Paul-Alexander Beckerburg. ISBN 978-3-384-16951-8
(Reprinted with kind permission of the author)



TOLERANT Software

2.2.1 Duplicates

Imagine standing in an infinitely large room full of identical doors with the same rooms behind them. This image may seem surreal, but it is an apt metaphor for one of the most complex problems in data management: the presence of duplicate data sets. Each of these doors – or duplicates – not only causes confusion, but also inefficient use of resources and potential errors in decision-making. But how do you navigate such a room to maintain an overview and ensure that every door you open leads to a unique and valuable space? In this text, we will take a closer look at the challenge and identify solutions that not only reduce the doors in our metaphorical room, but also ensure that every door opened leads to a room full of quality and value.

Data is more than just numbers and texts. It is the lifeblood of modern companies. This is where an innovative solution rises from the ashes of outdated data processing systems: TOLERANT Match. Imagine a world where incorrect customer addresses, duplicate records and inconsistent information are a thing of the past. This is exactly the world that TOLERANT Match makes possible.

The challenge of data quality in today's fast-paced business world is enormous. Companies collect vast amounts of data every day – from customer information to transaction data. But what happens if this data is incorrect, incomplete or out of date? The answer is simple: inefficient business processes, missed marketing opportunities and disappointed customers. But this is where TOLERANT Match comes in, a software that represents nothing less than a revolution in data processing.

TOLERANT Match goes beyond traditional data cleansing and offers a comprehensive solution for data quality. With advanced algorithms and an intuitive user interface, it enables organizations to efficiently cleanse, verify and enrich their data assets.

Whether it's identifying duplicate customer records, validating addresses or updating outdated information, TOLERANT Match does it with unparalleled precision.

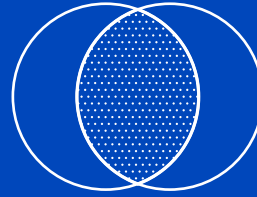




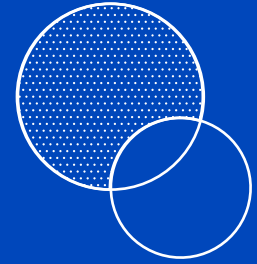
Find customer quickly with data entry



Keep your customer base free of duplicates



Merging data sets



Remove negative lists from stocks

Application areas TL Match

A core feature of TOLERANT Match is the error-tolerant search and matching. This technology makes it possible to find matches even with slightly different data entries – a game changer for companies working with large customer master data. Imagine being able to find precise matches despite minor data entry errors such as typos, different spellings or even missing information. TOLERANT Match makes this possible, improving not only data quality but also customer satisfaction and retention.

The flexibility and adaptability of TOLERANT Match is another key to its success. The software can be easily integrated into existing IT system landscapes and supports a wide range of data formats. This means that companies of any size and in any industry can reap the benefits of TOLERANT Match without having to overhaul their existing systems. In addition, the software offers configurable parameters that allow users to customize data processing to their specific needs.

Not forgetting the impressive performance and efficiency of TOLERANT Match. The software processes data with a speed and reliability that significantly increases the efficiency of business processes. From

processing speed to scalability, TOLERANT Match is designed to meet the needs of modern businesses.

At a time when the quality of data is directly linked to a company's success, TOLERANT Match offers a solution that not only overcomes the challenges of data processing, but also opens up new opportunities for growth and innovation. It is a powerful tool that helps companies unlock the full potential of their data and secure a competitive advantage in the digital era.

In the big picture, TOLERANT Match is not just a software solution; it is a promise for a future where clean, consistent and up-to-date data defines the difference between success and failure. For companies ready to take their data quality to the next level, TOLERANT Match is the key to realizing this goal.

2.2.2 Outdated information

Outdated information in databases is like ghosts from the past – it whispers of a reality that no longer exists. Imagine navigating through the narrow streets of a medieval town with a map that still shows the old city walls that have not existed for centuries. The situation is similar with outdated information in the modern business world: it leads us down blind alleys, leaves us standing in front of closed doors and, in the worst case, causes financial losses.

The dynamic landscape of the business world, in which companies merge, realign or close, customer data changes due to relocations or name changes and products are constantly evolving, requires equally dynamic data management. The challenge is not only to capture the constant change, but to map it in real time in the databases so that decisions can be made on the basis of up-to-date information.

The consequences of outdated information can be diverse and serious:

Marketing campaigns that come to nothing because the target group no longer lives at the specified address, increased operating costs due to redundant communication attempts or even the risk of running into regulatory difficulties due to non-compliance with legal requirements.

Against this backdrop, it becomes clear that updating databases cannot be a one-off affair, but must be a continuous process.

Technologies and strategies to combat outdated information include implementing real-time data reconciliation, using cloud services to ensure access to the latest data sources and establishing a cultural shift within the organization that emphasizes the importance of high-quality data. In addition, data enrichment services play a central role by enabling existing data sets to be regularly reconciled and updated with external sources.

Overcoming the challenge of outdated information requires a shift in thinking: from data management as a necessary evil to recognizing data maintenance as a critical factor for business success. This is the only way to ensure that the ghosts of the past no longer lead us astray.



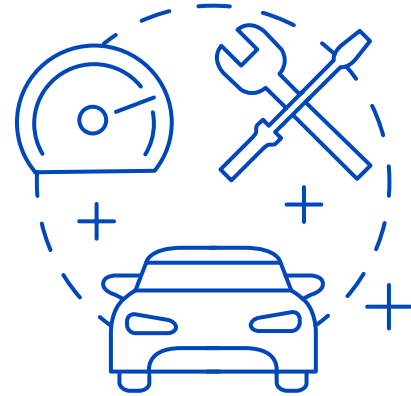
2.2.3 Inconsistent data formats

The issue of inconsistent data formats is a daily headache for many companies. It may seem inconspicuous at first glance, but the impact on the efficiency and effectiveness of business processes can be enormous. Imagine a global company receiving thousands of data records every day, collected from a wide variety of sources, each time in a different format. The variety ranges from date formats that are sometimes American (MM/DD/YYYY), sometimes European (DD.MM.YYYY), to currency details that are sometimes written with a comma and sometimes with a decimal point, to address formats that vary from country to country.

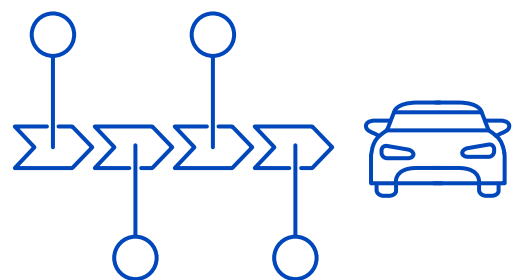
These discrepancies not only lead to considerable additional work in data integration and processing, but also harbor the risk of misinterpretation, which can have far-reaching consequences.

A misinterpreted date, for example, can lead to delays in supply chains, while an incorrectly converted currency can result in financial losses.

The challenge is to find a way not only to accept this variety of formats, but also to process them efficiently. The solution lies in implementing software solutions that are capable of automatically recognizing and converting data. Tools such as TOLERANT Match play a key role here. They offer the possibility of building a bridge between the worlds through intelligent algorithms and adaptability to different data formats. This ensures that data can be correctly integrated, analyzed and used regardless of its original format.

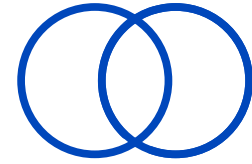


A clear example of this is the automotive industry, where parts and components from all over the world come together. An inconsistent date format in the supply chain can quickly lead to misunderstandings and delays. By using software that automatically detects and corrects such inconsistencies, such risks can be minimized. This not only leads to more efficient data processing, but also improves the quality of the data and therefore the basis for management decisions.



Overcoming the challenge of inconsistent data formats therefore requires not only a rethink in terms of data processing, but also a willingness to invest in technologies that make this diversity manageable. With the right tools and strategies, companies can significantly improve the quality of their data and thus gain a clear competitive advantage.

TOLERANT Match



Duplicate matching and data cleansing for address lists and customer data

Well-maintained address lists and customer master data are the basis for successful marketing and sales campaigns. In addition, an error-free address database is crucial for meaningful analyses and statistics. This is the only way to make operational and analytical decisions that will move your company forward. It is therefore important to avoid duplicates as early as possible in the data entry process, to carry out a regular duplicate check to find duplicates and to clean up the addresses. With TL Match, you can find these duplicates automatically, accurately and error-tolerantly.

Fast and error-tolerant search

TL Match finds address data, products and customers quickly and accurately, even in large databases and even if only fragmentary information is available. The special technology in TL Match also makes it possible to combine different search options.

Merge customer data from different sources in an error-tolerant, pure and clean way

Data protection legislation is making it increasingly difficult to use external data sources for customer data. This makes the high data quality of your own address data-bases all the more important. Addresses must be correct: Only if your customer data is clean can you

utilize and use it effectively. However, the volume of data is inevitably increasing. And customer data often has to be maintained and synchronized in different systems or databases in parallel. With TL Match, you can merge data from different sources easily and without duplicates.

Accurately find duplicates with TL Match

TL Match finds what belongs together. The easy-to-use software is accurate and error-tolerant. Powerful search algorithms find data records that match as soon as they are entered, but also when addresses are subsequently cleansed or matched from multiple sources. TL Match's error-tolerant search takes into account spelling or typing errors in names and addresses as well as different spellings. You can also adapt TL Match to your needs and requirements at any time. For the best results, clean data and duplicate-free address lists.

Some examples:

Dataset	Name	Street	Zip code	Location
Query	Alex Schepanski	Margretenstr. 9		Muenchen
Result	Alexander Sczepansky	Margretenstr. 7	80331	München
Query	Christina Maier	Bachstr. 1		Weissach
Result	Kristina Meyer	Bachstr. 1	71287	Weissach