## **InfoSec Advisory with Del Aden**

## **The business costs of dirty data: why clean & legally compliant data matters!**

## Most organizations struggle to make actionable data available, let alone turn it into business value. Research suggests only 10% of business-relevant data is actually used for analytics.

## This is because Business data can decay at a rate of over 40% per year, so your database needs to be kept up-to-date if it’s going to consistently deliver the results you expect. OVER TIME, YOUR BUSINESS DATA BECOMES INACCURATE.  Dirty data can be caused by a number of factors including duplicate records, incomplete or outdated data, and the improper parsing of record fields from disparate systems.



## The Data Warehousing Institute (TDWI) estimates that dirty data costs U.S. businesses more than US$600 billion each year**.** It is fair to assume there is a huge cost to companies and organisations in Africa as well!

## **But why do we have dirty data?**

## **Incomplete data** – missing phone numbers, postcodes, and other essential contact data can waste valuable time and lose customers.

## **Errors in data** – incorrect and misspelt contact data can damage your company’s image and affect the response from your customers.

## **Incorrectly formatted addresses** – careless data entry or incorrect loading of data can also damage your company’s image and get a poor response from your customers.

## **Obsolete data** – your customer data may have been correct at the time of entry but can decay with time. Out of date entries can result in high bounce rates and lead to a poor company image at a high cost.

## **Illegal data** – which can result in the risk of high financial penalties and potential prosecution if you are found to be holding illegal data

## **The impact of dirty data**

## Dirty data results in wasted resources, lost productivity, failed communication—both internal and external—and wasted marketing spending. In the US, it is estimated that 27% of revenue is wasted on inaccurate or incomplete customer and prospect data.

## Productivity is impacted in several important areas. Data scientists are spending around 60% of their time cleaning, normalizing and organizing data. In the meantime, knowledge workers are spending up to 50% of their time with hidden and inaccurate data.

## Dirty data lacks credibility, and that means that end-users who rely on that data spend extra time confirming its accuracy, further reducing speed and productivity. Introducing another manual process leads to more inaccuracies and mounting inconsistencies through growing numbers of dirty records.

## In addition to the revenue loss, dirty data impacts businesses more insidiously. Only 16% of business executives are confident in the accuracy that underlies their business decisions. Garbage in, garbage out—when you can’t rely on your own data, something needs to be done to increase data accuracy and reliability – i.e. Data Cleansing!

## **Cost of dirty data in banking & finance**

## Worldwide, inaccuracies in data costs between 15% and 25% of revenue for a company. With global revenues of over US$2.2 trillion, this means that dirty data costs the global banking industry over US$400 billion. Dirty data also leads to a number of risks that are unique to the banking industry. Inconsistent information across data silos in an organization leads to transactional risks such as inaccurate or even fraudulent transactions. Fake and fraudulent accounts should be caught early by processes that clean or detect dirty data. When they don’t, the bank is put at risk, and its reputation is damaged.

## **Dealing with dirty data**

## The most challenging problem in cleaning up dirty data is the cleaning of invalid entries and duplicate data. Careful error correction is needed to not only ensure that no data is lost while improving the consistency of existing valid data, but that all of the metadata corresponding to data correction is maintained alongside the integrated data itself.

## **Best practices for cleaning dirty data include the following practices:**

## Harmonize by correlating the data across different siloed sources and harnessing metadata for data provenance and lineage.

## Leverage core smart mastering capabilities to match and merge entities in a single multi-model platform.

## Apply semantics to capture relationships between data and to ensure consistency.

## Create a 360-degree view by integrating all of your data sources.

## Find dirty data using natural language searching, data modeling and machine learning to identify patterns and anomalies.

## **So, what are the benefits of data cleansing?**

## Reduces costs of your mailing

## Increases response rates to your campaigns

## Ensures your data complies with legal requirements

## Eliminates duplication and improves accuracy

## Gives you confidence when targeting prospective customers

## Consequently, every organisation must strive to increase the accuracy of their database in order to achieve the four basic principles of data security such as Confidentiality, Integrity, Availability and Authenticity.

## **In conclusion**

## **Activate your data cleansing project & accelerate business value**

## The business impact of dirty data is staggering, but an individual organization can avoid the morass. Modern techniques and technology can minimize the impact of dirty data. Clean, reliable data makes the business more agile and responsive while cutting down on wasted efforts by data scientists and knowledge workers.

## Your business might already be planning to tackle its dirty-data problems. However, you need to be aware that many data quality initiatives won’t address core changes needed inside the database to affect positive change where it is needed the most. This will only lead to future problems with inconsistent data, exacerbating the current state as data proliferates.

## In this regard, Delta3 International stands ready to help your business get the right start on cleaning its dirty data. We can help you improve data governance and increase the quality of your data assets overall.

## In conclusion, it must be noted that an organization that uses strong data governance in addition to data-cleansing practices can generate up to 70% more revenue., as such we urge all organisation (large and small) to make your Data Cleansing Project a TOP priority.

## ***About the Author***

## ***Del Aden is a UK-based Enterprise Solution Architect and an Industry-recognized Information Security Expert with over 20 years of hands-on experience in Consulting, Training, Public speaking, and Expert witness testimony. With expertise in Digital Transformation, Cyber Security, Data Governance and Business Continuity, Del Aden focuses on helping customers prevent security breaches, implement Digital Transformation and advice on Business Continuity Strategies and Exercises.***

## ***Contact:*** [***Deleaden@delta3.co***](mailto:Deleaden@delta3.co) ***| WhatsApp:+44 7973 623 624 | Web:*** [***www.delta3.co***](http://www.delta3.co)

