



infoboss

Best practice guide
to
Data Quality Management

info@infoboss.co.uk | www.infoboss.co.uk

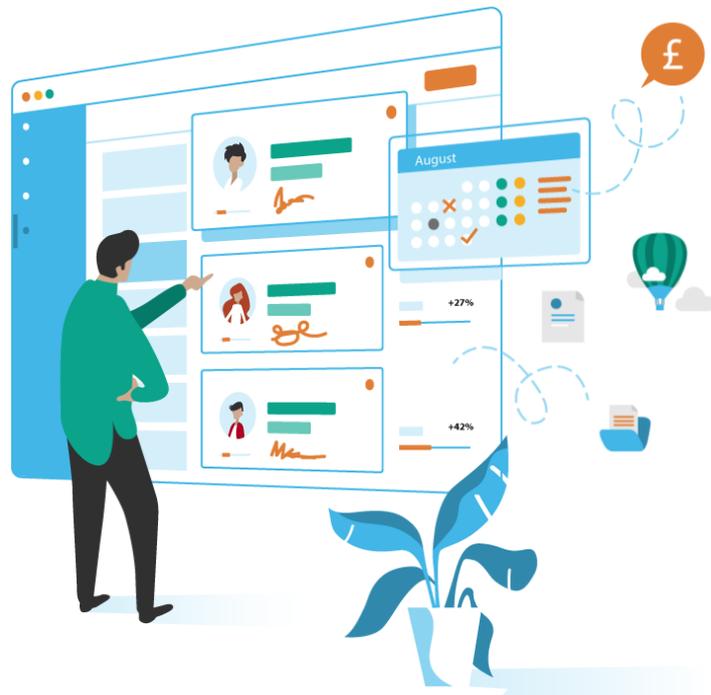


Table of Contents

1	Introduction – the business case	3
1.1	Cost of poor-quality data	3
1.2	Benefits of good quality data	3
1.3	The GDPR factor	4
1.4	Getting started... ..	4
1.5	Conclusion.....	4
2	Defining the problem, scoping and resourcing the project	5
2.1	Defining the problem, things to consider... ..	5
2.2	A simple process for improving data quality	5
2.3	Resourcing the project	6
3	Using software tools to support your Data Quality Management project	7
3.1	Building your data inventory	7
3.2	Data discovery.....	7
3.3	Action list extracts.....	7
3.4	Data quality project monitoring (performance indicators)	8
3.5	Data validation	8
3.6	Compliance with standards.....	8
3.7	Automated monitoring and audit	8
4	Conclusion.....	9
4.1	How Infoboss helps	9

**“The average cost to an organisation of poor-quality data in 2018 is \$15 million p.a.”
- Gartner (2018)**

1 Introduction – the business case

We've been in business for over 30 years, and almost every year we hear businesses say that poor-quality data is holding them back.

Fixing data quality is a daunting task and too often protagonists pay lip service to the job in hand and elect to adopt short term strategies to resolve specific issues rather than take the brave steps to build a data quality management strategy that delivers sustainable improvements, forever.

However, we believe this is beginning to change. With so many contemporary business drivers: digital transformation, artificial intelligence, automation, analytics and insight being underpinned by a deep reliance on quality and compliant data, businesses are being forced to confront the problem head on and take the necessary steps to embed a sustainable data quality management (DQM) or Data Governance (DG) methodology within the business operating model.

This paper explores the business opportunities presented by adopting a sustainable approach to improving data quality. The opportunity to embed a long-term strategic approach to data quality management is open to all business leaders, showcasing the business as a data leader, and delivering a genuine competitive advantage both today and in the future.

1.1 Cost of poor-quality data

Research from Gartner has shown a marked increase in the cost of poor-quality data on an organisation.

“The average cost to an organisation of poor-quality data in 2018 is \$15 million per annum!” – (Gartner 2018)

This is up from \$9.7 million dollars in 2017, and is a trend that is unlikely to improve due to the reliance of an increasing number of business initiatives, on quality data.

“Two-thirds of all business leaders believe that their companies must pick up the pace of digitalization to remain competitive.” (Gartner 2017)

The consequence of poor-quality data is therefore that initiatives that rely on data, such as digital transformation, and business analytics, will not deliver the business outcomes expected. In extremis they may fail completely if both the data quality is not improved, and the required processes and procedures for effective data input, collection and use are also not put in place.

1.2 Benefits of good quality data

By way of contrast, there are many benefits to having quality data beyond the obvious, including:

“The goal of any data quality strategy is to ensure that an organisation has and can use the data it needs to be successful”

- **Decision making:** Better quality data helps to inform decisions and lowers the risk of failure in delivering the expected outcomes. This improves business confidence that the decisions being made are correct.
- **Productivity:** Staff can be more productive as they don't need to spend time making sense of erroneous data.
- **Compliance:** GDPR and other industry regulations mean that you can no longer escape fines or reputational damage if you have data issues. You must have good quality data, and related data management practices in place, to support continued compliance.
- **Marketing:** Better data underpins advanced analytics and enables more focussed marketing efforts. It also facilitates omni-channel customer engagement and is the basis for a single customer view and an organisation's ability to differentiate themselves in a digital world.

1.3 The GDPR factor

GDPR has presented a unique opportunity for advocates of data quality improvement. Indeed, with the focus on data arising from this regulation, it is worth asking the question:

Has there ever been a better time to start a data quality improvement initiative?

Leading up to the 25th May 2018 deadline, GDPR compliance was top of many business to-dos lists. The focus was (for once) squarely focussed on data: Where is it kept? What are we storing? How and why are we collecting it? Is it secure? Who is responsible for it? How are we going to service data portability and subject access requests? And many more data centric questions. Indeed, all questions that a good data quality improvement project will also want answers to.



Moreover, GDPR has driven leading organisations to look at the GDPR's "Data Protection by Design and Default" mantra as an opportunity to review their operating model to ensure continued GDPR compliance, and to be more efficient in their data management practices. This presents a great opportunity to combine a review of data management practices with a data quality initiative and kill two birds with one stone.

1.4 Getting started...

Daunting is often a word used by those responsible for the data quality improvement initiative as they contemplate what's ahead. However, clearly articulating the business case is a great starting point. To help you do this, there's a great article published by Gartner on [developing the business case for data quality improvement](#).

1.5 Conclusion

The time has never been better for advocates of data quality improvement within the enterprise to come to the fore and make the fundamental changes necessary to data collection, processing and use by putting in place a long-term sustainable approach to improving data quality. Doing so will greatly increase your business competitiveness and chances of success with your current and future business transformation initiatives.

2 Defining the problem, scoping and resourcing the project

In this section we delve a little deeper into the processes involved in improving data quality, and how you might scope and resource the project.

2.1 Defining the problem, things to consider...

Experts generally point to four areas to consider for your “improving data quality” initiative.

- **Accuracy:** Data must be correct! How can you post a letter to someone if the address is wrong?
- **Completeness:** Have you collected all the data you need to? For example, how can you support customers in a certain age band if the date of birth is not being collected or defaults to 1/1/1900 – all customers are over 100 years old!
- **Standardised:** Particularly relevant when seeking to share, analyse or compare data with others. If you’ve only collected half of what you need for the standard view, then how can you process or share it effectively?
- **Authoritative:** Can you rely on the source of the data? What is its data lineage? Does it have credible inputs? Is it fit for purpose? This applies whether the source is internal, or external, to the organisation.

2.2 A simple process for improving data quality

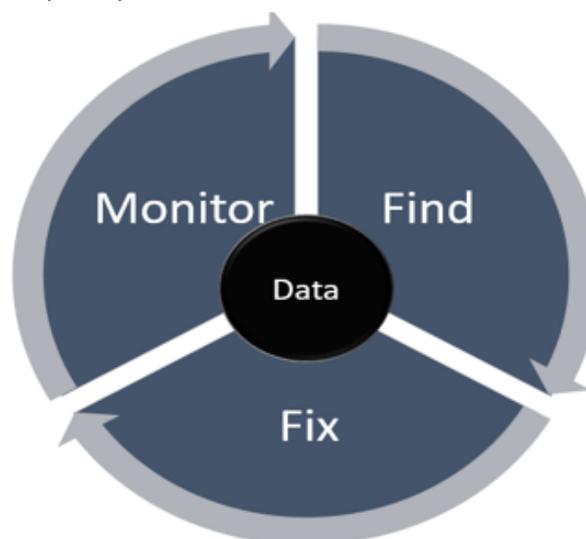
There are three simple stages to improving data quality:

- 1) Find it;
- 2) Fix it (at source); and
- 3) Monitor it (to ensure it is being collected and processed correctly in the future).

Finding poor quality data is perhaps easier to do now (post GDPR deadline) than it was, as many organisations have undertaken data audits to understand where data is stored and how it flows through their organisation. Importantly they have also taken the first steps in assigning data owners (or stewards) to take responsibility for the data within their sphere of influence.

Of course, this is where software can begin to play a part. It can be extremely difficult to discover poor quality data without both the tools to do so, and on the scale needed for the modern enterprise.

Modern data management tools such as Infoboss allow you to collect data from source systems, classify it, assign data owners, and then apply easy-to-use data discovery techniques to run rules over your data. In so doing, this will expose areas for improvement and allow you to baseline the problem. This will help you to not only understand the scale of the problem, but, importantly, to begin to identify the measures that you will subsequently use to monitor progress towards your data quality goals.



Powered by


Fixing data necessarily needs to be done at source. Having identified where data quality issues are apparent, these need to be fixed within the source system. In tandem with data correction it is important to identify, and fix, the root cause, be that by correcting the data collection process through staff training, or by enhancing application code in source systems to provide better validation and control on data input.



Monitoring your data is often overlooked. As a result, many initiatives to improve data quality subsequently fail and need to be repeated, just to contain the problem. No matter how robust your systems and processes are, there are inevitably going to be data collection activities undertaken by humans and these can, and will, lead to error. We read recently that a data quality improvement initiative tracked the level of introduced data quality issues by their own project team. This was a staggering 3% of records! And this was a team setup

to fix the data!

The hard reality is that fixing data is not a one-off activity. It's for this reason that data monitoring is crucial to support a sustainable data quality management culture by enabling data owners to continuously strive for excellence.

Tools such as Infoboss can apply machine learning techniques to constantly monitor the data stream through rules and alerts to data owners. Progress can be tracked by a series of configured dashboards for monitoring data quality performance measures.

2.3 Resourcing the project

So how do you resource your data quality management project for success?

- 1) **Appoint a project sponsor.** This should be a senior member of the management team with the authority to make decisions across the business that may affect resources, processes and systems. Remember you're looking to fix data quality once and for all, so half measures won't cut it! This person also needs to be aware of any changes to strategy that might impact the success of the initiative – like a merger or acquisition. A good candidate for this role would be your CIO or Head of Operations.
- 2) **Appoint a project manager.** A direct report to the CIO, this person will be responsible for managing the delivery of the project. This person will probably come from the IT team. Remember that IT systems and changes to processes surrounding them are likely to play a significant part in the project.
- 3) **Establish a centre of data quality management excellence.** This would be a small group of people from across the business that can advise others and get their hands dirty if necessary. Likely to be subject matter experts, data analysts, business analysts and good quality administrators - all having a can-do attitude.
- 4) **Finally, appoint a team of data owners (or stewards).** These will have authority over process, people and systems to not only fix data, but monitor it into the future and ensure it stays fixed.



3 Using software tools to support your Data Quality Management project

Having built the business case, scoped and resourced your project, the final step is to find the tools to empower your project team to efficiently deliver a successful outcome. Of course, if your organisation only uses one or two data collection systems, it may not be practical, or make commercial sense, to implement a system to help your initiative. However, most organisations have more systems in play than they would like to admit to. These systems feed other systems and the data estate grows exponentially as it works its way into the management information and decision-making processes of the business. Poor data governance surrounding these data collection and “value add” systems are the breeding grounds of poor-quality data. What’s needed to support your project is a consistent approach to firstly explore your data’s quality and baseline the problem across all types of data source, and then secondly a system that can constantly monitor your data for quality degradation in the future.



Although not an exhaustive list, the below are some of the practical features to look for in a DQM tool.

3.1 Building your data inventory

A good starting point will be to leverage the good work you’ve done for GDPR compliance when conducting your data audits. You will have undoubtedly collected a list of data processes and sources of data used within the business and perhaps even identified owners of these data sources.

Your DQM tool will need to be able to not only record information about your data sources, but also to stream in the relevant data to a central repository enabling data quality analysis to take place. Given the potential volume of data that you might have, it will necessarily need to be a highly scalable solution to house it all.

3.2 Data discovery

Once the data has been inventoried you will need to collect and store it in a format ready for rapid analysis. Once loaded you will want a simple interface to search and filter your data to gain insight into the quality of it. For example, you may wish to identify all instances of inaccurate data, missing data and non-standard data. Perhaps, you may want to understand the percentage of quality data to total ratios, run validation rules on fields such as dates, postcodes, phone numbers, currency amounts, salutations etc. Essentially you need to build a picture of the scale of the problem.

3.3 Action list extracts

Having identified the data that needs fixing you will need to be able to supply the list of problem data records to the data owners. The ability to easily extract data into work lists and possibly even supplement the lists with corrected values in order that the source systems can be updated, is an

essential requirement It may even be useful to include links against each data record that take you back to the source system to correct the erroneous data without having to extract an action list.

3.4 Data quality project monitoring (performance indicators)

You will want to monitor the progress of your project towards its goal. For example, ensure 95% of records have a valid postcode, mobile number, date of birth or person gender. Your DQM tool should allow you to setup the measures you want to monitor progress towards your data quality improvement goals.

3.5 Data validation

DQM tools should enable you to establish data validation rules. For example, validation of a payment card number, phone number, date, postcode etc. Also, custom validations that not only check the format but will lookup the value against an authoritative list to ensure data validity. You will want this functionality to be extended to any field that can be verified against a master list, for example, counties, countries, towns, NHS numbers, bank sort codes, suppliers, customer references etc.

3.6 Compliance with standards

Where data is required to be captured to meet a data standard, then validation and completeness rules will be needed to check that a data record meets that standard. You will want to easily setup these rules and put them in place for continuous monitoring of your data.

3.7 Automated monitoring and audit

Whilst people are involved in the data collection process, there is an inevitability that things will go wrong again in the future. A crucial component of your DQM tool is the ability to establish data quality rules that will run against your data estate forever. The tool should be automatically monitoring and checking the data across the estate to find exceptions and alert data owners as they occur and before they become a problem.



4 Conclusion

Data Quality Management (DQM) are the practices and procedures your business puts in place to ensure the delivery of high-quality information to all the consumers of that information, be they customers, stakeholders, employees or agents. DQM covers all aspects of data processing from collection through analysis and distribution. Your business will not be able to make accurate and informed decisions, deliver on its strategic goals or mitigate against service delivery issues or compliance non-conformances if the data being used throughout the business is of poor-quality.

In this paper we have provided insight into the business case, project scope, resourcing and requirements of the tools you will need to put in place a sustainable DQM strategy. Hopefully, it has provided inspiration and motivation to commence your data quality improvement initiative.

4.1 How Infoboss helps

Our team have expertise in the challenges posed by poor-quality data and how to correct it. We recognise that correcting the quality of data is not a point in time event. Its for this reason that we designed and built the Infoboss DQM solution.

Infoboss enables you to get your DQM project up and running quickly, delivering quick wins, finding data of interest, generating work packages but also putting in place long term automated processes to:

- Identify poor quality data;
- Alert data owners as inaccurate or incomplete data enters the data estate;
- Validate and suggest data correction;
- Track your data quality improvement performance indicators; and in turn
- Build a sustainable data quality improvement culture within the business.

We support the principal of democratisation. Our tools empower those that can make a difference to data quality, providing them with the tools needed to identify and resolve issues as they arise.

If you're looking for DQM tools or personnel to support your data quality improvement initiative we'd be keen to talk.

Please get in touch via email info@infoboss.co.uk or via our web site www.infoboss.co.uk.