

DATA QUALITY & INTEGRATION

Integrate data quality into your data projects: **5 foolproof expert tips**

Business lines, managers, CIOs, data stewards, integrators:
integrating data quality is easier than you think!





INTRODUCTION

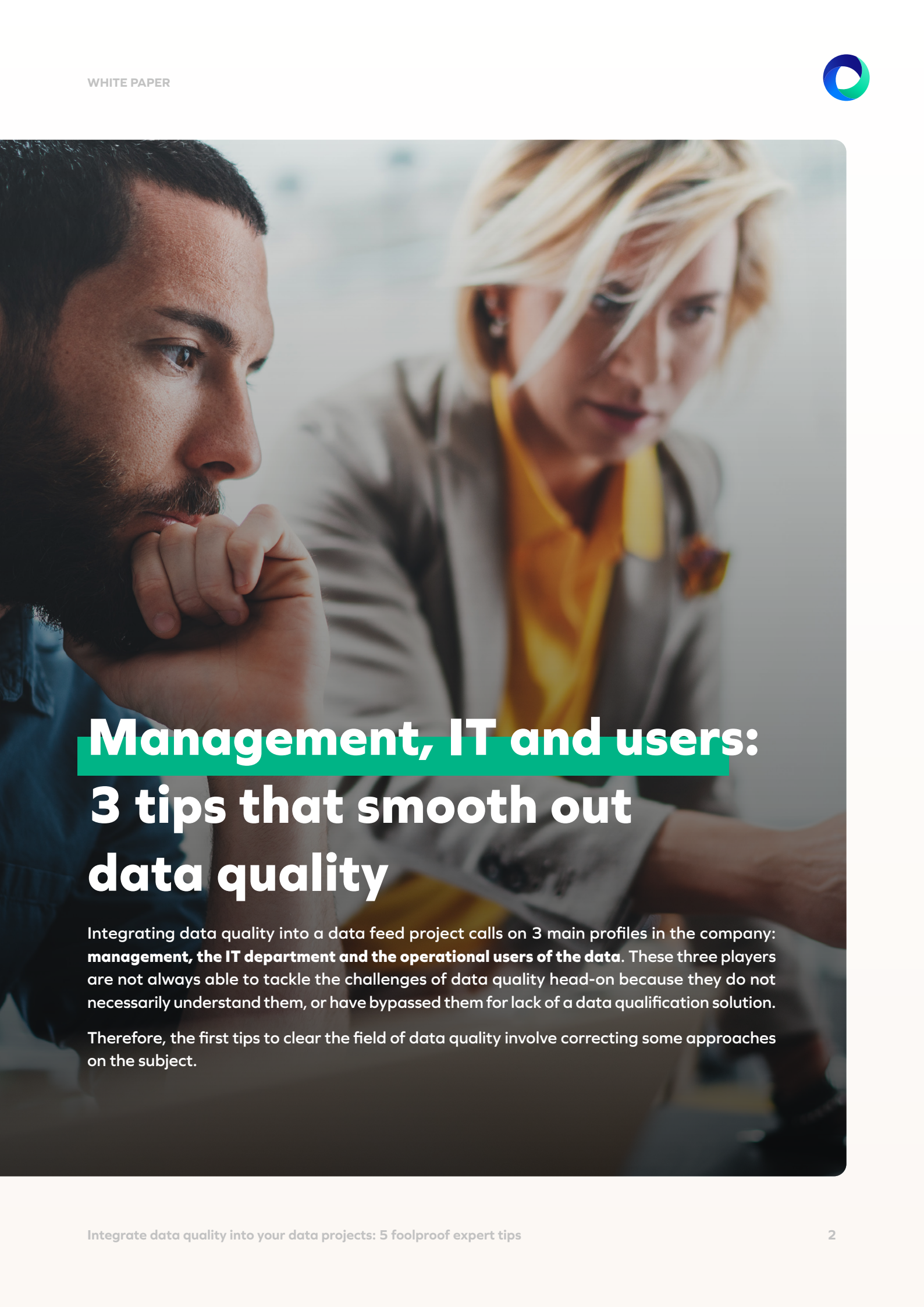
DP, CRM, UCR, Marketing Automation... As soon as a project concerns customer data, data quality comes to mind, as well as the problems associated with it.

When data quality is handled later, it can be a painful experience, and must be dealt with in a hurry, without resources associated with “a project that is not actually a project”.

This is not inevitable: data quality can be anticipated and integrated without difficulty.

This white paper gives 5 tips - plus a bonus - to stop suffering from data quality as an integration difficulty, and to make it a lever for optimizing your data projects.





Management, IT and users: 3 tips that smooth out data quality

Integrating data quality into a data feed project calls on 3 main profiles in the company: **management, the IT department and the operational users of the data**. These three players are not always able to tackle the challenges of data quality head-on because they do not necessarily understand them, or have bypassed them for lack of a data qualification solution.

Therefore, the first tips to clear the field of data quality involve correcting some approaches on the subject.



MANAGEMENT, IT & USERS

Management does not see the problem

In its steering role, management is not aware of the operational details and hassles. As long as operations are running and reporting metrics are not showing suspicious numbers or variations, it is difficult to pinpoint data quality issues. However, these distort numbers in reporting.

A case in point is a client whose objective was to increase its customer base. After a year of using the CRM and creating new customers, our tools revealed that salespeople were creating duplicates to benefit from the incentive. And this was happening in spite of the firewalls deployed by IT (single email address for example). The reporting to the steering committee suggested a 25% annual increase in the base. However, the management was in fact faced with distorted reporting, in a context with data that was itself distorted by duplicates, but without any warning signs since there was no impact on turnover.

TIP N°1

Carry out a quality audit of the database

This type of audit is an effective indicator of data quality problems. Contact information that is incorrect or that cannot lead to successfully reaching the intended target, presence of duplicates. Even scanning just a sample of the database provides a quality audit report that identifies data quality issues in the systems. Management then gains visibility by accurately identifying:

- the share of duplicates in the database,
- the share of deliverable emails,
- the number of invalid or unreachable telephone numbers,
- the share of undeliverable postal addresses, etc.

**27%**

of executives do not know how much of their data is valid.

SalesIntel, 2022



Taking care of your data is the best way to avoid strategic problems in the short and medium term, and to increase the value of data that is healthy and searchable for future needs.

Olivier Thierry,
co-founder and CTO of
mobile.club



MANAGEMENT, IT & USERS

The IT department does not bear its share of responsibility

Without a dedicated data quality team, data users tend to turn to IT when faced with data quality issues in customer data feed projects. But IT departments do not think it is their job to deal with this issue. Indeed, the mission of IT remains above all to ensure the technical infrastructure capable of circulating data flows. On the other hand, intervening on the content of these flows does not fall within the scope of its mission. In other words, once the technical answer has been provided, if the data flows entering the system are of poor quality, it is primarily the responsibility of the business lines and data managers. And when the project of deploying a data quality solution comes up, the main concern of the IT department is to avoid heavy integration, an increase in complexity of the IS, or redevelopment. And all this, without sacrificing the capabilities and advanced settings to finely customize data quality to the company's uses.

**70%**

of companies do not have a dedicated data quality team.

Study by O'Reilly, 2020

TIP N°2

Opt for a data quality solution with rapid technical integration

Among the technical integration facilities to look for in a solution:

- connectors with the main business line tools on the market,
- ready-to-use APIs,
- minimal need to recode, and if necessary, easy-to-use code,
- fast and proven deployment procedures.

A data quality solution that brings together these technical integration facilities wins over the IT department's support, and allows it to bring real added value on its scope.



In a data quality project, it is important to synchronize the capabilities and not underestimate the load on internal tools like ERP. Another tip: coordinate well with the various stakeholders in a global approach to the project.

Gai Marom,
External CIO of IZAC



MANAGEMENT, IT & USERS

Business lines adapt to the lack of data quality

Operational staff members are on the front line when it comes to recognizing the weaknesses in the quality of customer data in the database. For example, when faced with duplicates: a customer service department gets into the habit of juggling several records when calling the same customer. This symptom of duplication in the database considerably slows down operations, hence the implementation of a makeshift solution to get around the obstacle.

It is not uncommon to see departments writing micro-procedures, sometimes on a case-by-case basis, to explain where to find the right information for a customer among their various records. Operational staff members adapt to these kinds of practices, however time-consuming and unproductive they may be.

These micro-procedures are far from solving the problem, are quickly lost when an employee leaves, and are used as a palliative solution in the absence of a real data quality solution and can hide more serious problems in the long run.

TIP N°3

Give control to the business lines, and free them from having to rely on IT, who will appreciate it!

Operational staff members know better than anyone what data quality problems need to be eradicated. The stakes are high, since installing data quality boosts their operational performance and avoids bypass strategies.

A must-have: a data quality solution that allows operational staff members to fine-tune in complete autonomy, and that is available directly from the interface of their daily tools such as CRM. For this reason, you should avoid 'black box' solutions!

**+ 80%**

of data users
experience data
quality problems.

Datafold, The State of Data
Quality in 2021



A data quality project requires the sharing of information with the business lines concerned with the customer data. When you know what you want, you progress faster!

Michael Bouyer,
Project manager in
charge of website
security and data
protection, Bip&Go

A man with a beard and glasses is looking at a bar chart on a screen. The chart shows a steady increase in values from January to December. The background is dark and blurred, focusing on the man and the chart.

Data stewards & integrators: 2 tips for anticipation

As strategic stakeholders in a data feed project, data stewards and integrators have an essential role to play in terms of anticipation and raising awareness.



DATA STEWARDS & INTEGRATORS

Data stewards manage the scramble around data quality

Between the functional logic of business lines and the technical approach of IT, data managers have the skills to reconcile the two worlds. They are able to identify data quality issues, but they also need to foster cooperation between the different players who still operate in silos in most companies.

Fostering cooperation for strategies of both sides in terms of data quality can be an impossible equation, without a real specialized solution dedicated to customer data quality.

TIP N°4

Promote data quality with a specialist solution

Deploying a data quality solution helps to eliminate bad practices, for example by blocking form validation when there is data entered that is inaccurate, invalid or already in the database.

In addition, data quality helps acculturate teams to customer data with a convincing argument: the results it brings in terms of operational efficiency. But only if you deploy a solution designed by a specialist.

Indeed, generalized solutions or worse, homemade solutions, are to be avoided, because they give very uncertain or even false results and can lead to believing that the problem is solved.

Note that tip #3 also concerns data stewards: a solution that gives them control gives them the power to act!

**45%**

of data users want to improve data reliability and quality.

Datafold, The State of Data Quality in 2021



Thanks to the integration of DQE, the results obtained in emailing demonstrate the interest of working on the quality of the database, more than on its volume. These results help renew the marketing culture in the company through qualitative work on data.

Marion Camurat,
Head of Marketing,
Zaion



DATA STEWARDS & INTEGRATORS

Integrators want to perform an acceptance test of their tools without data quality problems

Many integrators have had the painful experience of data quality problems that mar the acceptance testing of a large customer data project, just as the operational staff members are getting to grips with their new solution. When data quality has not been addressed upstream, it disrupts operations at this stage. Inability to reach a customer because of an invalid phone number, a database with a sudden increase in the number of contacts (due to emerging duplicates): faced with these anomalies, business lines complain that the project does not provide the expected benefits. The delivery can be called into question and cost extra hours of work to find out how to rectify the quality of the data, future billing disputes between stakeholders, etc.

TIP N°5

Anticipate data quality with a trusted publisher

There is no inevitability. Integrators must have the reflex of including data quality in all projects that are in contact with customer data and of proposing a proven solution that can be easily integrated into a larger project. In addition to a full functional scope and state-of-the-art capabilities, in order to identify the right publisher, the specification must include:

- experience and legitimacy with a strong offer in terms of support,
- certification program, guaranteeing a solid partnership,
- support for customers in the deployment and use of the solution.



On average,
45%
of leads lack qualified data,
33%
are duplicates.

Integrate report: "It's not you, it's your data"



We wanted to work on the quality of our customer data above all, aware that without data quality, a marketing tool will not give the expected ROI.

Ludivine Jallet,
CIO of Pacific Pêche



THE SPECIAL DEDUPLICATION TIP

Know where to stop!

Data quality remains a complex subject, especially in terms of deduplication of duplicate data. Also, when a data quality solution gives good results, companies tend to want to 'push' it to the limit for maximum results. However, data quality is always about finding the best compromise.

It is impossible to identify 100% of the duplicates in a database.

Thus, our algorithms allow for the calculation of a percentage of proximity between the duplicates.

100% proximity duplicates are certain, while records with 50% proximity will certainly not be duplicates. In practice, we have found that the ideal proximity percentage is between 80% and 90%, which allows for automatic merger processing.

Generally, the duplicates that remain in the database become negligible and, above all, they can be rectified on a case-by-case basis as the company's practices are refined, without consuming precious user time.



The ideal proximity percentage is between

80% & 90%

When it comes to data quality and deduplication, you also have to know where to draw the line!

This good practice contributes to a successful integration of data quality with objectives that are realistic, and that are reached. Defining this type of KPI requires expertise in data quality, which is why a project in this field must imperatively include a specialist. Their role is to provide proven benchmarks to ensure the success of your data quality project.





TO SUM UP

5 tips to easily integrate data quality

TIP N°1

Conduct a quality audit of the database

Strong visibility for management of the quality of its information database thanks to the audit report on data quality problems in the systems.

TIP N°2

Opt for a data quality solution with rapid technical integration

Simple technical procedures reassure the IT department and allow it to provide real added value to its scope.

TIP N°3

Give control to the business lines and free them from having to rely on IT, who will appreciate it!

Operational staff members know best what data quality issues need to be addressed, and they need to be free of workaround strategies that degrade business line user performance.

TIP N°4

Promote data quality with a specialist solution

A specialist solution helps data stewards reconcile the functional expectations of the business lines with the technical constraints of IT.

TIP N°5

Anticipate data quality with a trusted publisher

For integrators, it is the assurance of avoiding dysfunctions related to data quality in the acceptance testing stage of their large data projects.

★ BONUS TIP

Deduplication: Knowing where to stop

Understanding that it is not possible to identify all duplicates in a database is a definite step forward in the success of your data quality project!



About DQE

Founded in France in 2008, DQE Software (DQE) is the vendor of DQE One, a modular data quality management solution.

Our solutions are specialized in BtoC and BtoB customer contact and identity data. DQE One controls the quality of data including last name, first name, e-mail address, telephone number, postal address, and legal data. DQE One includes two product lines: DataQ, a contact data quality solution, and Unify, a deduplication solution.

Across all industries, DQE helps organizations ensure that their customer data is reliable, and accessible to everyone, at all times. DQE One stands out as the most powerful data quality solution on the market, capable of managing databases of several tens of millions of contacts. Each year, our processing represents over 3 billion queries with an average response time of 150 milliseconds.

DQE One relies on powerful engines and algorithms that benefit from over 10 years of experience. The solution interfaces with more than 240 international databases for the postal address portion. Natively, DQE One offers connectors with solutions such as Salesforce, Microsoft/Dynamics, Cegid, Magento, etc. Our algorithms work on all our customers' clouds: Heroku, Azure, AWS, and GCP.

DQE One is used by a wide range of business lines (CRM teams, Chief Data Officers, Chief Digital Officers, CIOs, customer relationship managers) for more than 400 clients. Among them are several listed companies and well-known brands such as Groupama, EDF, BUT and Belambra.

For more information, visit www.dqe.tech/en and follow us on   