DATA QUALITY AS A KEY SUCCESS FACTOR FOR MIGRATION PROJECTS

(Practice Oriented)

Dephine Clement
A.I.D.
dclement@aid.fr

Soumaya Ben Hassine Guetari
A.I.D.
sbenhassine@aid.fr

Brigitte Laboisse
A.I.D.
blaboisse@aid.fr

Abstract: Organizations are facing the challenges of data migration, which results either from mergers and acquisitions due to the consolidation of the economy, or from upgrade of systems or finally from willingness to simplify the data storage architecture. A data migration is not a trivial task; either, it is a significant project undertaking. Because the data stored in the legacy systems is a strategic company asset, it needs to be analysed, measured, preserved and improved before being brought over to the target system. In this paper, we present business cases and best practices on how companies structure their data quality management in the particular context of data migration. We propose a comparison grid with classified items such as the project management teams, the data quality audit and the key quality indicators (KQIs), the data quality improvement activities, the optimum quality level calculation, the problem of the data references and the need for a data governance post migration.

Key Words: Data Migration, Data Quality, Project Management Teams, Change Management, Users' Survey, Data Governance, Value of Data, Key Quality Indicators.

1 INTRODUCTION

Due to business consolidation, mergers and acquisitions, companies working more on a cross-functional basis and less in silos, there is a need to simplify the data storage systems architecture, to create reference data and centralize them in a unique repository. The business need for data migration exists today. But is

there research on the aspects of project and data management in the context of migration? It seems that the answer is “not really”.
We started by looking at the ICIQ proceedings and at the most popular topics [11], and we did not find any paper related specifically to migration or merger and acquisitions; we found papers around data quality in the context of data warehouses[12], we found papers around multiple sources of data and data quality at the source[13], papers about master data management roles to put in place[4], the IQ indicators to measure[2], but generally speaking, we did not locate any academic paper proposing a data quality/data management framework in the context of a data migration project.
And yet, it is critical to take into account all Data Quality aspects in a data migration effort. To illustrate, we would like to cite an article from Data Migration Pro [21], published in August 2009, which is saying: “There needs to be a complete DQ framework in place from day 1 of the migration otherwise every stage of the project is open to risk, delay or failure.”Another article [14] published on eWeek by Brian Prince in May 2008 was using the striking title of: “Don’t let data quality concerns stop a data migration project”.

In addition, in the industrial world, migration projects and mergers/acquisitions often give the opportunity to create awareness around the importance of data quality and to put in place a data quality strategy. The following questions are emerging during a migration :

- What are the roles and responsibilities of the DQ team versus the Data Migration team ?
- How should the unique referential be built; should it allow local references; which tradeoffs; how should it be controlled ?
- Should the business rules be standardized before migration?

Multiple white papers are published by software vendors, data migration specialists or research companies such as the Microsoft CRM Data migration Framework [8], the Utopia data migration methodology [22] and the Bloor-Research article [5] for example; those whitepapers provide either a simple guideline [10] or a more sophisticated method such as Velocity from Informatica [17].

These frameworks are the result of business cases and a collection of best practices. Because it collects business experiences and it is feedback sharing, it is very helpful in terms of guidelines, but it is often biased by the software used. For instance, a big part of Velocity is related to the Powercenter implementation (Informatica ETL and data quality tool).

2 BACKGROUND

A French Energy Production Company (later referred to as FEPC) counts 12 active Production Plants around France. The data around the plant industrial assets is currently stored in information systems that are heterogeneous from one plant to another: different customization of the initial same technology, different business rules and different usage of the data.
In order to standardize and rationalize the architecture of the industrial asset information systems and to be able to do consolidated strategic reporting, the FEPC has decided to migrate over a unique “off the shelf” Enterprise Asset Management (EAM) Platform.
This effort is currently in the preparation phase. The migration actually starts in 2011 and is supposed to end in 2015.
The volume of industrial asset data to be migrated over to the EAM is over a million records; in addition to the data, asset maintenance processes are also to be migrated.
The project management team is composed of 12 people from the FEPC and is supported by its R&D department, as far as the data quality work is concerned. This R&D department appointed A.I.D. to execute the data quality work: from the data users interviews and the data auditing, to the recommendations and actual data corrections.

What is interesting to note is that the FEPC migration project team appointed the R&D department to assist on the DQ piece of the work, which we can interpret in the following way :

1) Data Quality was on the radar screen of the migration project team, which is a good news for our domain!
2) The "evangelization" on the importance of DQ, started three years ago by the R&D department through conferences, proofs of value and papers is paying off
3) Data Quality remains a relatively "new" subject, which in this case requires the expertise from the R&D department

3 PURPOSE OF THIS PAPER

The purpose of this paper is to focus on the roles and responsibilities of a data quality team in a migration project, to highlight the key DQ related topics relevant to a data migration, and to illustrate the "modus operandi" through three business cases; the purpose is NOT to propose a methodology for data migration.

First, we describe the context of the company FEPC, the migration project, the questions and hypothesis. Those initial recommendations were made as a starting point to the FEPC.
Second, we introduce 3 business cases, focused on product migration and similar to the FEPC context. Those business cases aim at validating our initial hypothesis and assumptions. The interview of the 3 companies was conducted face to face for Company B and Company C and over the phone for Company A. It was conducted in May and June 2010.
Third, we benchmark what we observed with the 3 business cases and our initial approach in terms of the roles and responsibilities of the DQ team and the key DQ topics for a data migration.
The conclusion highlights the key success factors, the major DQ key points to take into account when starting a data migration journey. This helped A.I.D.'s recommendation to FEPC.

The FEPC asked for recommendations in 5 main areas, as detailed below. For most of the questions, our answer to FEPC was 'YES' based on good practices in DQ management , even if we did not have a formal, theoretical support as explained above. Because the repercussions in terms of resource consumption and project planning were high, it was necessary to validate the hypothesis and convince the project team through the confrontation with similar business cases.

- H1 : Is it necessary to create a unique industrial asset referential ?

Impact : FEPC has a list of materials per plant. To create a model of industrial material, it means de-duplication intra-plant and inter-plant and resolution of divergences. The process is complex because, for very different spelling names , the model from one plant to another can be the same. We can also find the same name for different materials, the comparison requiring extra attributes not directly available.
Our initial recommendation : yes, it is necessary to move to a unique referential, in order to optimize the spend across the various material vendors and to standardize the industrial asset maintenance plan (frequency, duration, profile, …), which is critical from a safety management standpoint.

- H2 : Is it necessary during migration to standardize the business rules ?

Our initial recommendation : yes, in order to do cross-plant consolidated reporting.

- H3 : Is it necessary to clean the data before migration ?

Impact : In terms of project management, the initial option retained by FEPC was to migrate the data, map it into the new model, without preventive cleansing and to schedule a cleaning project post migration. Rearranging the order of the cleansing steps meant for FEPC a significant change in terms of planning.

Our initial recommendation : clean the data before migration for 2 reasons : the migration can be the opportunity for the users to discover the data under a new light and find out new anomalies. The second reason is from a cost perspective : cleansing the data after migration is typically more costly because the new system is not necessarily designed to allow cleaning process such as de-duplication and merge.

- H4 : It is recommended to track and filter out obsolete records before migration ?

Our initial recommendation: yes, in order to avoid to clean obsolete data during the migration and to limit the final system in term of volume, reducing the maintenance cost.

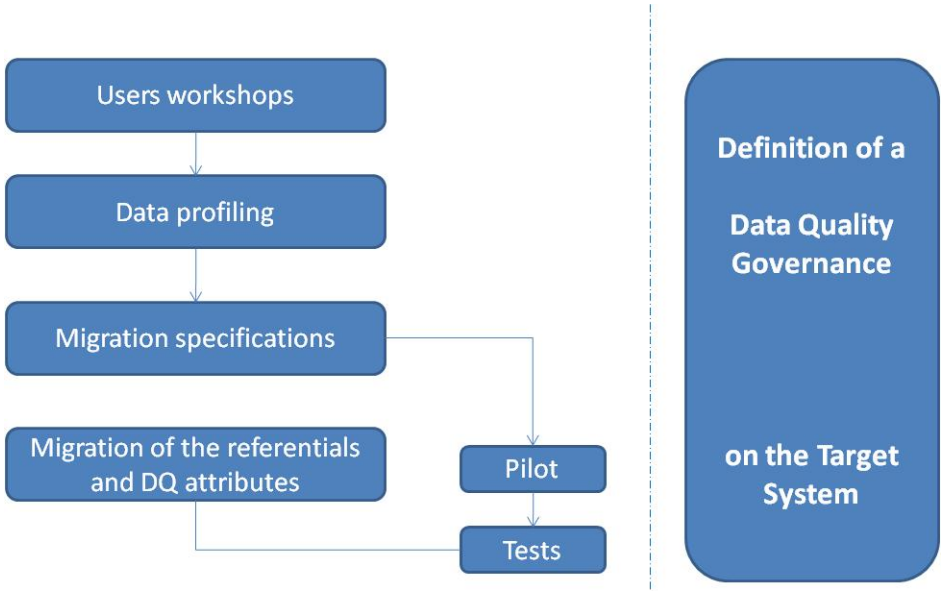
- H5 : Is it necessary to create a DQ team ? Which profile? To do what?

First recommendation from A.I.D. to the FEPC project team: The role of a data quality team in a pre-migration mode should be the one highlighted in section below.

3.1 The Data quality team : our first recommandation

3.1.1 Role and responsibilities

The main tasks of a DQ team within a data migration project are summarized on the figure 2 below:



DQ Team : Main tasks - Figure 1

3.1.1.1 Users workshops

The objective of these workshops is to collect the users' perception on the actual DQ level (subjective point of view), to understand the DQ dimensions which are important in business/ financial terms, to assess the data « value », to agree on an optimum data quality level. This part is not specific to a migration project and is relevant for each DQ project [18] [23], [7].

3.1.1.2 Data profiling

During this phase, the metrics were calculated according to the data specifications given by the users.
This part is not specific to a data migration project either.

3.1.1.3 Migration specifications

We distinguish in the data model 2 kinds of entities, transactional and referential, [20]:

- Transactional Entities : purchase orders for instance
- Referential Entities : products for instance

The DQ team is in general more focused on the Referential Entities part because the data manipulations call methods that are well known in the Data Quality domain:

- De-duplication or customer/product identification to detect the same products between different sources with similar but not equal names and/or codes [175
- Obsolescence detection [6], [16] : to filter obsolete data, to avoid to pollute the new system with, for instances, no more existing products.
- Multi-sources management [3],[9] : to consolidate attributes (for instance the color, price,..) related to the products data overlap found.

On the transactional part, the DQ team's role is about documenting the business rules in place, to measure the level of compliancy to those rules and to detect and value the anomalies.

Based on the DQ findings from the analysis of the Referential or Transactional Entities and based on the output of the users' interviews, the DQ team will be in a position to write the data migration specifications, as far as the level of effort and the priorities of data cleansing are concerned.

3.1.1.4 Pilot and tests

Within a « Data migration» project, the role played by the DQ team is part of the overall data mapping work that needs to occur between the legacy systems and the target system.
Depending on the size of the DQ team, the availability, the selected tools, it can play an active operational/technical job or "just" a role of consulting / specifications /pilot / tests.

3.1.1.5 Definition of a DQ governance on the target system

Data Quality should not be limited to a one-off data cleansing project with the objective to make the data compliant for the migration. The migration is an opportunity to put in place a true data governance strategy including aspects such as data management processes, organization responsible for the data quality and data quality measures. Putting in place such a strategy implies:

- Defining the role of each teams/individuals vis-à-vis the data: data creators, data custodians, data consumers [1].
- Documenting data management processes as far as data entry, data maintenance and data administration are concerned
- Publishing regularly data quality measures and alerting if below the green target (optimum quality level defined during the users' interviews)

As it is well known, data quality is a continuous improvement cycle [19]; it is therefore beneficial to capitalize on the learning and the experience gathered from the data migration and to put in place a robust well defined data governance strategy.

3.1.2 Profile of the DQ team members

The DQ team should include people with a double profile: information system and statistics. Migrating data from legacy systems to a unique target central system often requires a “DQ toolkit”, a project management approach that is at the boundary of data modeling and statistics techniques.

4 Best practices sharing with 3 companies that undertook the same projects

In order for A.I.D. to validate the initial assumptions and make accurate recommendations to the FEPC R&D department and Migration Project Team, we organized best practices sharing session with 3 international companies having undertook or currently still going through a company wide data migration. Those 3 companies were chosen based on the following factors:
- Mandatory : they completed or started a data migration project
- Mandatory : the migration was from « n » information systems to one target : the goal is to have the case of matching and merging across multiple sources, not ‘only’ one single system to migrate.
- Optional : the data migrated was material or product
- Optional : material maintenance processes were to be migrated as well

For confidentiality purposes, we will refer to them as Company A, Company B and Company C.

Company A is a US based company doing business in Europe as well. This company operates under different brands and in 2005, it decided to have all functions from all brands running from a unique SAP ERP platform. The overall migration project is planned to end in 2011. The data migrated is client but also product data.

Company B is a French based company doing business all around Europe. In 2001, it decided to standardize the product offering of one particular business branch. This effort was aiming at better managing stock, at improving inter-company selling, at better serving European clients and finally at providing a consolidated reporting on the product offering and the sales. To meet this objective, the company decided to migrate its product data over to a SAP unique referential repository.

Company C is a multinational company that underwent two big waves of B2B client data migration over to a unique Worldwide CRM system. It is still regularly undergoing smaller client data integration projects resulting from Merger and Acquisition activities. Even though the initial trigger of the 2 migration waves was political with strategic alliances to other companies, there was the intent to rationalize the client information systems architecture, to be able to uniquely identify the customer and finally to be able to do consolidated reporting and get as close as possible to a 360°C view of the customer.

Here is a table summarizing the similarities and differences in the data migration context between the 3 companies that we interviewed and the FEPC.

	Company A	Company B	Company C	FEPC
The referential data to be migrated is product data	Yes, product but also client	Yes, product only	No, B2B client	Yes, material
An industrial model of material data (or product hierarchy) was to be built	No, already there	No, Product offering already standardized	No, client data	Yes
Migration of "n" systems towards a unique target	Yes, multiple industry sectors	Yes, multiple countries	Yes, multiple countries, company acquisitions	Yes, one per plant
Material maintenance processes to migrate?	No	No	No	Yes
Number of "input" sources	Around 20	Around 15	Around 25	4 to 5 if we consider one plant = one source
Duration of the migration phases	16 to 18 months for the first system	1 to 1 year and a half	1 to 1 year and a half	Forecast is a year

Selection of companies for best practices sharing session – Grid 1

4.1 The Comparison Framework

For conducting the best practices sharing sessions with the 3 companies, and based on the needs and main questions from the FEPC Migration Project Team, A.I.D. grouped the items of interest in the following categories:

- A - The business objective of the data migration – reason for choosing to move to a unique data repository
- B - The various project management teams, their roles, profile of their members and responsibilities: data quality team, change management team, release management team
- C - The key data quality indicators used to measure the data before, during and after the data migration (referred to as KQIs)
- D - The type of data quality auditing and corrective actions prior to the data migration
- E - The management of data references: pros and cons of central versus local references

- F - The data quality toolkit
- G – The milestones of the project
- H - The implementation of a data governance to continuously monitor and improve the data post migration

Below is a summary of the answers of the 3 companies to each of the item, as well as the key learning for the section. This summary/comparison will allow us to make the proper recommendations to the FEPC.

4.1.1 Section A: The business objective of the data migration – reason for choosing to move to a unique data repository

For all 3 companies, the data migration projects are triggered because they serve 6 main business objectives such as:

- decrease of the number of information systems and simplification of their architecture : for Company C for example, the number of customer information systems WW before migration was 800. Post migration and consolidation, it went down to 200 for the entire customer information ecosystem. In addition, for each business function, **one unique** master reference was nominated and the architecture of upstream/downstream information systems was redesigned around this master reference. With 800 information systems, it was impossible for company C to govern the customer information and to avoid data quality issues caused by data propagation or error of synchronization
- serving and knowing international customers : for company C, where the business case was on B2B customer data, it is important to set up a proper WW hierarchy of the accounts (customers), to operate internal sales compensation and global customer knowledge management. In addition, the selling model is also dependent on the existence and quality of this global hierarchy. International organizations can have indeed central procurement teams which buy centrally for all subsidiaries. The WW unique referential of customers is therefore in such case very much strategic. This argument does not apply to a B2C environment.
- improving stock management to better control the procurement process and decrease the number of suppliers
- standardizing the product offering to increase sales efficiency and reduce the errors in the order management process
- consolidated reporting : for company B for example, post-migration, they were able to monitor the selling of products across subsidiaries and to further improve the convergence of local product references towards the unique product referential

4.1.2 Section B: The various project management teams, their roles, profile of their members and responsibilities: data quality team, change management team, release management team

The key learning of this section is that the 3 companies interviewed did create a **data quality team** from day 1 of the migration project. For company A and B, the data quality team was part of the migration functional team overall. Whereas for company C, it was a dedicated team functioning autonomously but of course working hand in hand with the other migration functional stakeholders. For all 3, the data quality team is a central team. However, this central team is supported by local business relays. For company B and C, the business relays are dedicated full time to the migration whereas for company A,

they are only working at 20% of their time on the migration project but all migration items are a priority to them.
The profile of the team members vary from one company to the other; for company A and B, the team is a mix of business and IT profiles, for company C, the profile is mainly business and the IT team is a satellite team supporting the DQ team. In company C, the DQ team is supported by an offshore execution centre of about a 100 people.
The breath of roles and responsibilities of the DQ team in the migration effort vary from one company to the other:

- For company A, the role of the team is mainly to collect the "data specifications" from the user, to document them for the IT migration team, and to test the corrections applied to the data according to the data specifications
- For company B, the role of the team is mainly to drive the convergence of the local product references towards the unique product reference: spot the obsolete references, correct and standardize surviving ones, rename and retag others.
- For company C, the role of the team goes from data measurement prior to migration, after data correction and after migration; assessment of the level of effort for the offshore team; recommendations of data cleansing based on data specifications; supervision of the data cleansing effort; testing that the corrections are in line with the data specs and mapping and standardization of the data management processes (preventive and corrective) so that they fit the new target system. This last step prepares for the data governance strategy post migration.

Finally, it is interesting to note the importance of **the change management team** – this aspect was taken into account for all 3 companies. For company A and C, it is co-located with the release management team. There is an interaction between the release management team and the data quality team as far as the interpretation of the business rules is concerned. Company C also told us about the strategy of the first "country" to migrate – for this company, it was important to select the most flexible country, the one the least reluctant to change.

4.1.3 Section C: The key data quality indicators used to measure the data before, during and after the data migration (referred to as KQIs)

The 3 companies interviewed are all measuring their data quality and have developed the concept of an optimum quality level that they want to reach before migrating the data over to the target system.
Company A is the only one to have blocked and delayed migration because this optimum level was not reached.
Company B and C are allowing for a buffer; they let records with a non optimum data quality level go into the target system; they flag them and correct post migration. However, as already stated above, correcting data once migrated may turn out to be much more difficult and costly. Indeed, for example, the target system may not be equipped with a smart merge tool and if the migration team lets too many duplicates go in, the correction of the duplicates post-migration may turn out to be very manual and thus resource-intensive.
Before deciding on which KQI's to put in place, all 3 companies starting by interviewing the users and gathering the data specifications. They then audited the data to understand what the "departure" data quality level was. Based on both, as well as on a risk/value study, they could properly recommend the optimum quality level to achieve for the data, prior to migration.
To be noted that the optimum quality level and the KQI's to measure will be reused in the post migration data governance mode.

4.1.4 Section D: The type of data quality auditing and corrective actions prior to the data migration

For all 3 companies, the data cleansing effort is very well structured. It always starts with a data users survey to understand what are the business processes supported by the data and what are the business rules managing the data – data specifications phase. It then continues with the audit of the data to understand its current level of quality and the level of effort to get to the agreed optimum level of quality. All subsequent cleansing efforts are based on the output of this round of users' survey and data audit.
3 main types of correction are being done:

- Detection of obsolete data. Company C flags the obsolete client records but for legal purposes (contracts attached to the client), it migrates those records over to the unique referential system. Those records are "blocked" in the target system and flagged as "obsolete" so that they do not propagate in the downstream applications.
- De-duplication of local referential data to create a unique referential.
- Standardization of business rules driving the data.

4.1.5 Section E: The management of data references: pros and cons of central versus local references

Company B and Company C have clearly documented the role of the central team versus local business teams in terms of the management of referential data. They have agreed on trade off selection criteria to decide if a reference should be central or kept local or made obsolete. The criteria are based on:

- Business justification and ROI
- Maintaining the consistency of the unique referential

4.1.6 Section F: The data quality toolkit

All 3 companies are using more or less sophisticated tools to help them in the data quality tasks. Company A and C, due to the high volume of data to migrate and the need to industrialize the data quality management processes, felt the need to be equipped from an off the shelve data profiling, cleansing and monitoring tool. Company B, because the volume of data was low and the tasks highly manual, was ok with using excel. So we can conclude that the choices for a tool are driven by the business needs, the budget constraints, the repeatability of the tasks to execute and the volume of data to manage.

4.1.7 Section G : The milestones of the project

Company B was the only one which structured its project around milestones. They build a real repeatable framework that works for all countries to migrate. The 4 high level steps are:

1) Preparation of the deployment: resources evaluation, cost analysis, scope definition, sign off from senior management
2) Business convergence phase: the output is a local referential data convergence to the unique referential
3) Execution of the convergence plan
4) Post go live support

Company C is starting to build up a similar framework for managing migration projects. This became a necessity with all the mergers and acquisitions that the Company is currently going through.

Company B and Company C were not willing to share in more details this framework as they consider it to be their industrial property and therefore under confidentiality.

4.1.8 Section H: The implementation of a data governance to continuously monitor and improve the data post migration

For all 3 companies, the DQ work started during the migration phase pursued with the implementation of a true data quality journey. DQ rules are documented and communicated, a DQ team is in place with a clear mandate (RACI chart, responsible, accountable, consulted, informed), DQ is monitored either daily or monthly and data creators are made accountable for the errors done. Senior management sponsorship is recognized as a key success factor of the data governance [17] strategy.
The companies capitalized on the data quality management knowledge that they acquired during the pre-migration phase.

5 COULD WE VALIDATE OUR HYPOTHESIS?

At the end of this round of interviews, we could validate our hypothesis and initial recommendations that we made to FEPC and that we presented in section 3 of this document.

Here is in short the key points that need to be highlighted:

- H1 : Is it necessary to create a unique industrial asset referential ?

As explained in section 4.1.1, the 3 companies opted for the creation of a unique referential (client or industrial asset depending on the context). The company B demonstrated the business benefits of managing such a unique referential; benefits in terms of reporting across subsidiaries and across countries and in terms of stock management. Company A and B explained how they managed the x-referencing (legacy reference, new reference); Company B even did a physical relabeling of its stock to conform with the new unique reference.

- H2 : It is necessary during migration to standardize the business rules.

Company A and B undertook that effort on the target system but also on some priority records from the legacy systems.

- H3 : Is it necessary to clean the data before to migrate ?

Regarding the correction of the data pre-migration, all 3 companies audited the data, analysed the level of effort and agreed on an optimum level of quality. This created their data specifications document and it drove the data cleansing effort.

- H4 : It is recommended to track and filter out obsolete records before migration.

Only Company B did not migrate the obsolete data; the other two were constrained to migrate them into the target system but “flagged” them. Company B tried to ‘keep alive” the legacy system in

parallel of the new one but the adoption of the new system was delayed by this decision and triggered change management issues.

- H5: Regarding the role of a data quality team in a migration project, only Company C had a dedicated independent DQ team. Company A and B bundled the DQ component with other functional aspects in a broader team. However, in data governance mode, post-migration, all 3 companies capitalized on the knowledge acquired pre-migration and re-use the same DQ team structure, this time dedicated for all.

6 THE KEY LEARNINGS

When we compiled and analysed the interviews, several success factors became very obvious as they were repeatedly stressed by the 3 companies as 'key elements'.

1) The use of a **project management framework** in order to structure and accelerate the project

Company B made it clear that the data migration project is done in phases and articulates around milestones. This company demonstrated the importance of creating, formalizing and reusing a project management framework. For each phase, the company assigned number and profile of resources, productivity and timelines. The framework allows for repeatability and acceleration of the various waves of migration. In the case of FEPC, the same framework would be executed for each plant's migration.

2) The strong **sponsorship of the senior management** as far as the data quality strategy is concerned

This item seemed "a given" to A.I.D. but was continuously stressed by the 3 companies as a key point to keep in mind. The sponsorship of the senior management and the fact that they strategically position the data migration project, will drive 3 main benefits: decrease the resistance to change, improve the cooperation of the business with the data migration team and make data quality a priority.

3) The network of **local business relays**

In order to understand the business needs and the business rules as far as the data is concerned, it is necessary to officially and publically create a network of local business relays. They are the warrant of the meaning of the data migration project. Of course, it is also necessary to free up their working time in order for them to dedicate to the data specifications work.

4) The use of **Key Quality Indicators to pilot** the project

Those KQIs are designed and agreed upon by the data quality unit within the migration project team and with the local business relays. Driving the project with the KQIs allows determining the level of effort to bring the data to its optimum quality level (green target) or within the agreed area of acceptance (yellow target). If the KQIs are green or yellow, this gives a "go" to migrate the data over to the unique target system. Those operational KQIs are also aggregated for project management review.

5) The **shutdown of the legacy system** on day 1 of the switch to the new system

As said before, Company B tried to maintain both systems alive in parallel, which increased the resistance to change and created data synchronization issues. The recommendation is therefore to completely shut down the legacy systems on day 1 of the migration and to block the read/write accesses.

6) An appropriate **data quality toolkit**, based on budget constraints, volume of data to clean and migrate and repeatability of the tasks to execute

7) The set up of a **data governance strategy** by capitalizing on the data quality knowledge acquired pre-migration

It is well known that data can only remain of good quality if it is continuously improved; indeed, the data, even if it is cleaned pre-migration, will naturally become obsolete and incorrect if not maintained in the target system. Therefore, all the pre-migration effort that consists in building a network of local business relays and a central data quality team, in driving the data quality by indicators and optimum level will be the baseline of the data governance strategy on the target system. The critical enterprise data asset will thereby be valued, protected and enriched.

7 CONCLUSION

As developed previously, our initial hypothesis as well as our best practices sharing session proved that it is critical to consider Data Quality as an important milestone of a data migration project. Data Quality is indeed a key success factor and an enabler of data migration projects. The data quality initiative pre-migration needs to be structured around several steps: assessment of the data needs and the business rules, definition of an optimum data quality level, audit of the data, data quality correction activities according to the data specifications; this strategy needs to be brought forward by proper DQ or data management teams with documented roles and responsibilities, and finally the DQ journey needs to be pursued post-migration through a data governance plan.

8 REFERENCES

[1] Burgess M., D*ata quality what is it, and do we agree?*, http://www.dmsg.bcs.org/web/images/stories/PDFs/where%20did%20it%20all%20go%20wrong%20-%20mikhaila%20burgess.pdf, 2007.

[2] Clément D., Laboisse B., Hewlett Packard Customer Master Program – Data Quality Audit of 6 Functional Masters, 11th International Conference on Information Quality, 2006

[3]Da Conceição M., Batista M., Salgado A.C. "Information quality measurement in data integration schemas". *VLDB'07*. 2007.

[4] Duff A., Master Data Management Roles, Their Part in Data Quality Implementation, 10th International Conference on Information Quality, 2005

[5] Howard, P. *Data migration: a white paper by Bloor Research*. http://whitepapers.zdnet.com/abstract.aspx?docid=322622, 2007.

[6] Hunter J., Choudhury S., *Semi-automated preservation and archival of scientific data using semantic grid services*, 5th international symposium on cluster computing and the grid, 2005.

[7]Lucas H.C., *A user-oriented approach to systems design*, 26th ACM annual conference, 1971.

[8] Manek, P. *Microsoft CRM data migration framework: white paper*, 2003.

[9] Mecella1 M., Scannapieco M., Virgillito A., Baldoni R., Catarci T., Batini C. "Managing Data Quality in Cooperative Information Systems". *LNCS 2519*, pp. 486–502, 2002.

[10] Mohanty, S. *Data migration strategy*, Information management special reports, http://www.information-management.com/specialreports/20040525/1003961-1.html , 2004

[11] Naumann, F. *Awards ceremony at ICIQ 14th, 2009*, http://www.hpi.uni-potsdam.de/fileadmin/hpi/FG_Naumann/ICIQ2009/ICIQ_2009_Awards.pdf, 2009.

[12] Neely P., A Proposed Framework for the Analysis of Source Data in a Data Warehouse, 6th International Conference on Information Quality, 2001

[13] Ponzio J. F., Jr., Authoritative Data Sources (ADS) Framework and ADS Maturity Model, 9th International Conference on Information Quality, 2004

[14] Prince B., Don't let Data Quality concerns stop a data migration project, http://www.eweek.com/c/a/Database/Dont-Let-Data-Quality-Concerns-Stop-Your-Data-Migration-Project/, 2008.

[15] Rahm E. and Do H.H., *Data cleaning: problems and current approaches*, IEEE Computer society technical committee on data engineering, 2000.

[16] Silvers F. “Building and maintaining a Data Warehouse”. 2008.
[17] Taylor, S. *Informatica Velocity: A proven methodology for integrating the enterprise*,, http://www.informatica.com/INFA_Resources/ds_velocity_6773.pdf, 2008.
[18] Wang R.Y., and Strong D.M. “Beyond Accuracy: What Data Quality Means to Data Consumers.” *Journal of Management Information Systems 12.* 1996. pp5–33.
[19] Weber K., Otto B., Osterle H., *One size does not fit all – A contigency Approach to Data Governance, ACM Journal of Data and Information Quality*, Vol.1, No.1, Article 4, 2009.
[20] Wolter R., Haselden K., The What, Why and How of Master Data Management, http://msdn.microsoft.com/en-us/library/bb190163.aspx, 2006.
[21] http://www.datamigrationpro.com/data-migration-articles/2009/8/19/some-recent-data-migration-lessons.html, 2009.
[22] http://ecohub.sdn.sap.com/irj/ecohub/go/portal/prtroot/docs/hub/uuid/e0d6974e-b4d6-2c10-d4aa-c06909bc611f%3FQuickLink=home
[23] http://www.legacy-systems.com/migration%20white%20paper.pdf

9 ANNEX 1 – DEFINITION OF ACRONYMS

FEPC: French Energy Production Company
EAM: Enterprise Asset Management
RACI: Responsible, Accountable, Consulted, Informed. We can sometimes even find the term RACIS where the S means Support
ERP: Enterprise Repository Platform
SAP: name of a German brand
B2B: business to business, enterprise customers
B2C : business to consumer
CRM: customer relationship management