



Belgium & Luxembourg

The impact of bad data on the Belgian grocery retail market

Mini Data Crunch Survey 2011

Survey carried out in collaboration with:



Connecting business
creating value



1. Introduction

1.1. Objectives of the study

Several GS1 member organisations in Europe (e.g. GS1 France and GS1 UK) have found out that data quality is still very poor in the grocery retail industry. Bad data have a severe cost impact on the industry as a whole as they have a very negative cost impact on the complete value chain of suppliers and retailers from ordering to pay. Also they have a direct impact on logistics (transport, warehousing, shelf capacity ...), sales and on shelf availability for the final consumer.

Given the results in other countries, GS1 Belgium & Luxembourg wanted to gain insights in the data quality situation between retailers and suppliers in the Belgian grocery market. The objectives of the study were set as follows:

- Study the current situation of the overall data quality;
- Identify the main problems around data quality at the retailers & suppliers side;
- Provide a first set of qualitative inputs for potential actions in the future.

1.2. Scope & methodology of the study

This study was conducted with **4 retailers** (Colruyt, Delhaize, Carrefour and Cora), representing more than 75% of the Belgian grocery market with the following selection criteria:

- at least 3 main retailers in the Belgian grocery market;
- some retailers active in the GS1 Belgium & Luxembourg data pool (CDB) and some not;

and **4 suppliers** (Unilever, Nestlé, L'Oréal and Alpro) with the following selection criteria:

- covering several product categories;
- some suppliers active in the GS1 Belgium & Luxembourg data pool (CDB) and some not;
- able to present at least 100 SKUs that are delivered to all 4 selected retailers.

"The exchange of required product data fields is expected to increase dramatically in the near future"

A quality data match was made of **34 product attributes** (24 mandatory data fields & 10 additional agreed data fields used in the GS1 Belgium & Luxembourg data pool - see Annex), at a certain reference date, between the ERP of each supplier and the ERP of the 4 retailers based on a **selection of 100 SKUs/supplier** that are sold at all retailers participating in the survey.

The 100 SKUs selected from each supplier covered:

- several product categories,
- regular and promotional items,
- only active codes (non obsolete),
- completely new products,

- products with attributes that changed recently.

1.3. Assumptions taken

To ensure the quality of the data matching study, a semi-automatic data matching was performed. Meaning that for instance different units of measurement were made equal e.g. 18,4 cm = 184 mm and a set of rules for rounding were applied.

If a **fully automatic data matching** had been performed, the results would have been **20% worse than with semi-automatic data matching!**

“Based on the sample of the survey, the ratio of mismatches reaches 50 to 70%!”



2. Findings data crunch

2.1. Data matching

The retailers' data mismatches comprise missing data and data which were not correct when compared to the suppliers' data (presuming the data of the suppliers are right). The sum of total

average missing and total average mismatches for the retailers ranges from 49% to 67% using the semi-automatic matching method (see Fig.1).

As for the suppliers, we did find some missing or wrong product data result-

ing in a total average missing + total average mismatches (=wrong product data) of 16% (see Fig.1).

It also became clear that the results turned for the worst when SKUs had a more complex product hierarchy.

Fig. 1	Missing data					Mismatches supplier - each retailer of available data					Total Missing + Mismatches				
	Sup.	Ret. 1	Ret. 2	Ret. 3	Ret. 4	Sup.	Ret. 1	Ret. 2	Ret. 3	Ret. 4	Sup.	Ret. 1	Ret. 2	Ret. 3	Ret. 4
Identification	9%	27%	15%	49%	48%	0%	9%	9%	9%	2%	9%	36%	23%	58%	50%
Dates	58%	67%	1%	77%	90%	17%	17%	57%	7%	5%	75%	84%	59%	84%	95%
Unit status	0%	47%	9%	99%	54%	0%	1%	70%	0%	30%	0%	47%	79%	99%	84%
Dimensions cu	13%	3%	1%	46%	44%	0%	18%	37%	28%	5%	13%	21%	38%	74%	49%
Palletization	5%	36%	1%	3%	8%	1%	18%	67%	45%	22%	5%	53%	68%	47%	30%
Content/weight	25%	67%	1%	6%	8%	0%	19%	64%	54%	36%	25%	87%	65%	60%	44%
Other info (Returnable, Country origin, Lifespan)	17%	29%	2%	9%	14%	1%	30%	65%	25%	33%	17%	59%	66%	34%	47%
Out of 24 mandatory attributes	14%	33%	8%	62%	56%	2%	10%	37%	11%	10%	16%	43%	45%	73%	66%
Out of 10 additional attributes	15%	40%	1%	6%	8%	1%	24%	72%	46%	32%	16%	63%	73%	52%	39%
TOTAL	14%	35%	6%	46%	42%	2%	14%	47%	21%	17%	16%	49%	53%	67%	59%

Having divided the 34 product attributes in 7 groups, the highest scores for mismatches + missing data for the retailers are to be found under the attribute groups:

- Dates;
- Unit Status;
- Palletization;
- Content/weight.

This has severe consequences. For example tracking of changes becomes impossible if 'dates' are missing and in

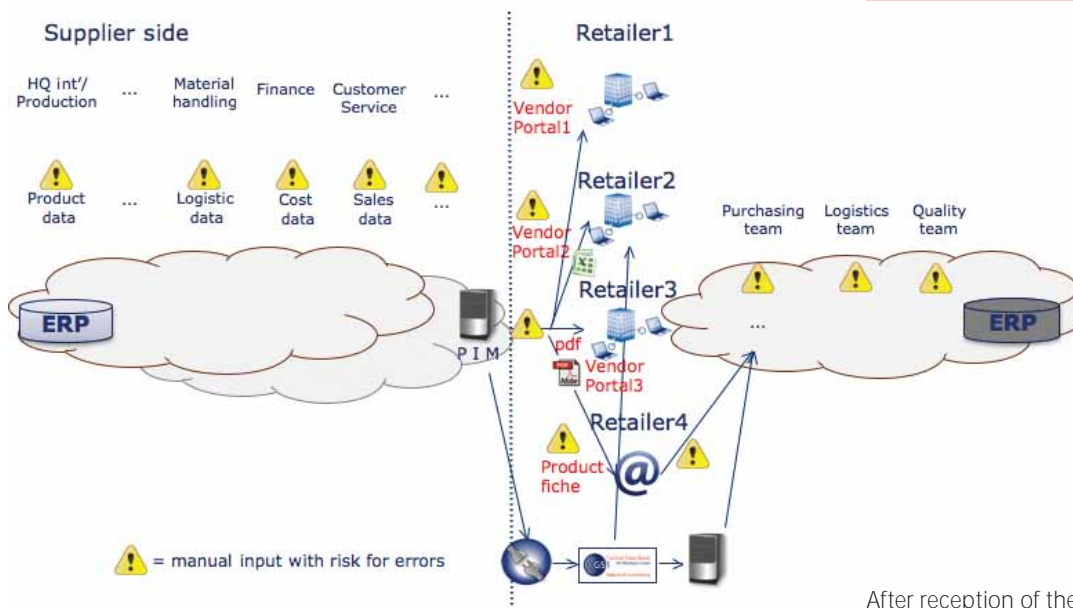
the absence of 'unit status' there is confusion between base unit & packs leading to ordering & invoicing errors. Also, warehousing & transport are affected negatively if 'palletization' and 'content/weight' are missing.

2.2. Information flow and IT architecture

During the study, the information flow and IT architecture at the suppliers and retailers was investigated more in detail (see Fig.2).



Fig. 2



Currently, on the supplier side, there is a first input of product data by the company HQ/Production. Different departments of a local subsidiary then add local, mostly logistic & commercial data. A final data quality check is done in most cases by a PIM (Product Information Management).

The data are then sent to the retail via:

- manual input in a vendor portal;
- manual input of a PDF sheet which is sent electronically and uploaded in the vendor portal;
- manual input of a product data sheet and sent by email;
- manual input of an Excel sheet which is sent electronically and uploaded in the vendor portal;
- or automatic data transfer if a data pool is used.

After reception of the data, from the retailers' side, there is:

- manual and/or (semi-) automatic check of the received data;
- manual input of additional product & commercial data for various purposes like pricing, channel management, logistics, store management (shelf labelling, planograms, ...), quality, ...

In other words, the information flow in general has many points of manual inputs increasing the risk of errors.

2.3. Cost Impact of bad data quality

Based on the same assumptions as used in the GS1 UK data crunch study, the Belgian grocery market will lose more than € 130 million within the next 5 years.

The costs involve 2 main areas:

- shrinkage costs: € 12,5 million/year
- workaround costs: € 6,2 million/year

Additionally, lost sales have to be taken into account adding at least € 7,9 million/year.

This brings us to a **TOTAL of € 26,7 million/year.**

Given this large cost impact, it is surprising that upon today, no operational dashboards on data quality at supplier neither retailer side exist and the impact of low data quality has not yet been quantified and translated into costs.

“We estimate the cost of bad data for the Belgian grocery retail market at more than 26 million per year if no action is taken”

800 attributes/SKU are managed in the respective ERP systems, but only 60 to 70 attributes/SKU are currently

exchanged. The attributes are both product related and commercial.

However, the exchange of required data (data fields per

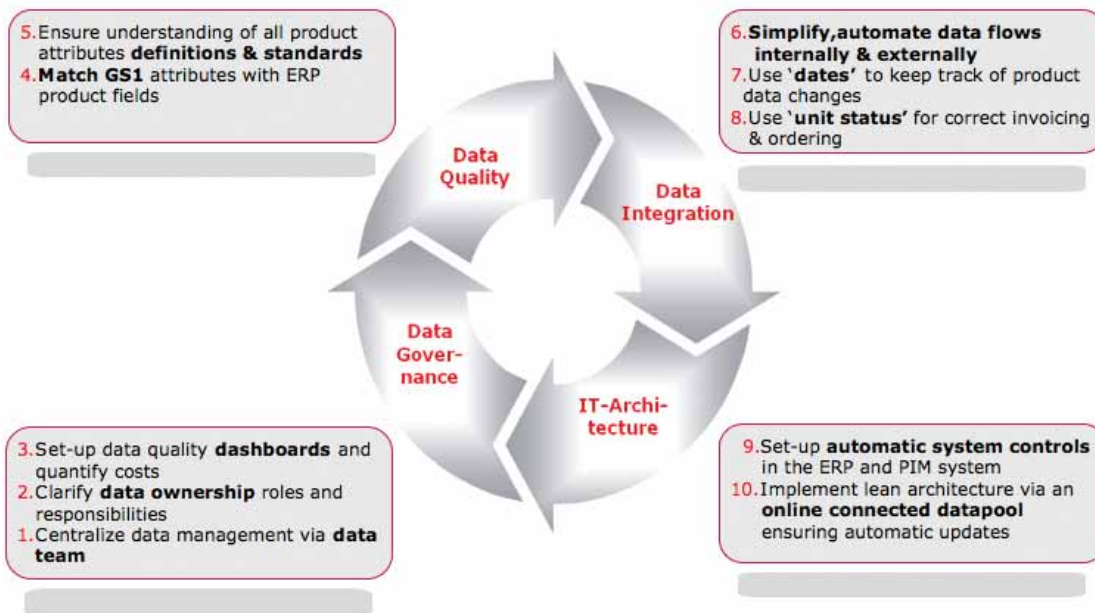
item) is expected to increase dramatically in the near future as many data fields will become mandatory to respond to environmental, health and other regulatory requirements. But also the final consumer is demanding more and better product information.

2.4. Attributes evolution

A qualitative survey among the study participants teaches us that on the supplier side about 600 attributes/SKU and on the retailer side between 300 and

3. Recommendations

BearingPoint Belgium, who conducted the study, makes the following recommendations to increase data quality:



4. Conclusions

Data Quality in the Belgian retail grocery market is very poor. Based on the sample of our survey, the ratio of mismatches reaches 50 to 70%. If we use fully automatic matching it approaches 80%.

Bad data are impacting the entire value chain from suppliers to retailers. As such bad data have severe cost implications. Based on the same assumptions as taken by the GS1 UK survey, we estimate the cost of bad data for the Belgian grocery retail market for € 130 Million if no action is taken.

The main cost categories include:

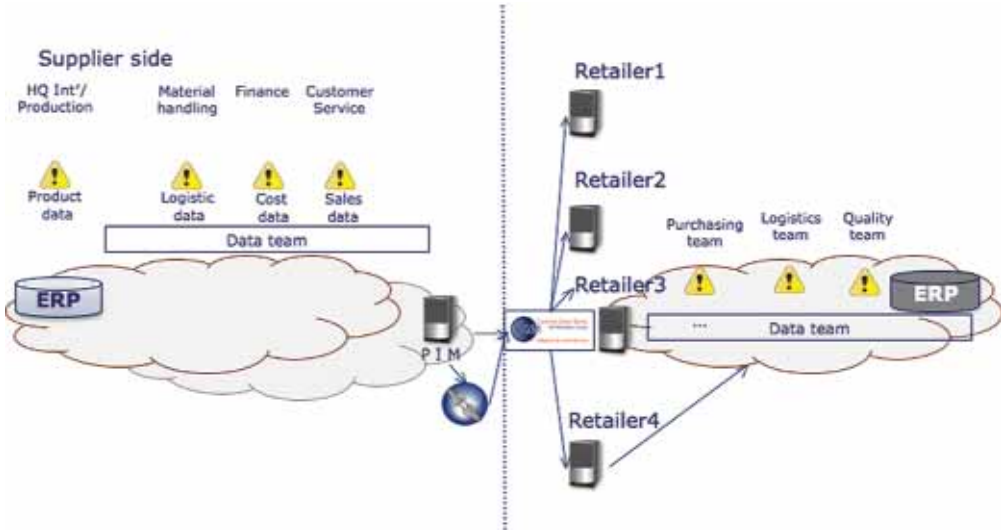
- shrinkage in the supply chain processes (47%) - eg. bad data leading to out of stock, wrong orders;
- workaround processes (23%) - eg. manual handling, matching invoices;
- lost sales (30%) - eg. perishable goods that stayed too long in the warehouse.

But solutions are available:

- centralized data management with single inputs of data managed by data teams (see Fig. 3);
- online sharing of a unique data source internally and externally which leads us to the use of data pools (GDSN).



An outlook on the future (Fig. 3):



In the scheme above, manual inputs at several departments are reduced by dedicated data teams at the side of the supplier and the retailer.

There is an automatic transfer of data put in place from the PIM of the supplier to the PIM of the retailer via e.g. the CDB, the GDSN data pool of GS1 Belgium & Luxembourg.

Automatic data quality checks are done at the level of the PIMs.

All this together will substantially reduce the need for manual inputs and the associated risk of errors.

5. Annex: Selected product attributes/SKU & grouping of attributes

1	Global trade item nummer	MANDATORY
2	Target market country code	
3	Trade item unit descriptor	
4	Effective date	
5	Publication date	
6	Classification catalogue code	
7	Information provider	
8	Name of information provider	
9	Brand name	
10	Start availability date	
11	Functional name	
12	Is trade item a consumer unit	
13	Is trade item a base unit	
14	Is trade item an ordable unit	
15	Is trade item a dispatch unit	
16	Is trade item an invoice unit	
17	Is trade item a variable unit	
18	Height	ADDITIONAL
19	Height UOM	
20	Width	
21	Width UOM	
22	Depth	
23	Depth UOM	
24	Is packaging marked returnable	
38	Country of Origin	
40	next lower level components	
41	Quantity of Trade Items per Pallet	
42	Quantity of Layers Per Pallet	
43	Quantity of Trade Items per Pallet Layer	
53	Net Content + UOM	
54	Gross Weight + UOM	
55	Net Weight + UOM	
66	Pallet Type Code	
73	Minimum Trade Item Lifespan From Time of Arrival	

- Identification (1,2,3,6,7,8,9,11)
- Dates (4,5,10)
- Unit status (12,13,14,15,16,17)
- Dimensions (18,19,20,21,22,23)
- Other info (24,38,74)
- Palletization (40,41,42,43, 66)
- Content/weight (53,54,55)

“The online sharing of a unique data source (GDSN data pools) will substantially reduce the need for manual input and the associated risk of errors”



6. Next steps

The most important conclusion from this study is the urgent need for datasynchronization. GS1 GDSN is a worldwide network of local certified data pools allowing continuous synchronization of product data amongst trading partners. The CDB (Central Data Bank) is the local GDSN data pool developed by GS1 Belgium & Luxembourg and operated by SA2 Worldsynchron.

Companies - member of GS1 Belgium & Luxembourg – can register to become a CDB user. They will be charged a limited annual fee for this service. They will benefit from a range of services such as:

Helpdesk

From the start, users receive the support of the multi-lingual GDSN team of GS1 Belgium & Luxembourg. The first level helpdesk is provided by a dedicated team of the organization, based on a ticketing system.

Trainings

GS1 Belgium & Luxembourg organizes conferences, collective and individual trainings. The organization also provides an e-Learn platform and webinars.

Solution providers

Our website contains an online tool allowing to find solution providers which can support you in your projects. See www.gs1belu.org/nl/zoek-solution-provider or www.gs1belu.org/fr/recherche-fournisseur-de-solutions.

Committees and working parties

The GDSN committee decides the general strategy lines of the CDB, whilst working groups discuss specific (technical) issues of the system. All CDB users can apply for participation.

More information?



Belgium & Luxembourg

Koningsstraat 29
1000 BRUSSEL

T +32(0)2.229.18.80
F +32(0)2.217.43.47
E info@gs1belu.org
www.gs1belu.org



Our helpdesk

CDB.support@gs1belu.org

GDSN/CDB website

www.gs1belu.org/fr/gs1-gdsn-0 or www.gs1belu.org/nl/gs1-gdsn

With thanks to the participating companies:

