

Data Quality GS1

The importance of data quality in the food industry

Arno van Rijswijk, Sander Swartjes and Ruurd van der Ham

In order to improve the data quality of product data in the food industry, the non-profit GS1 organization created a whole new solution, including third-party Data Management Services that will administer or check the product data from suppliers on quality. This must solve many challenges which are faced in this area and ensure that regulations are complied with.



A. van Rijswijk is a data quality and analytics manager at GS1 Netherlands. arno.vanrijswijk@gs1.nl



S. Swartjes is a manager at KPMG IT Advisory, Enterprise Data Management. swartjes.sander@kpmg.nl



R.F. van der Ham is a senior consultant at KPMG Enterprise Data Management. vanderham.ruurd@kpmg.nl

Importance of data quality in the food industry

High quality data is increasingly important for answering the demand from consumers in the food industry. Inaccurate or incomplete product information can lead to complaints from customers and even non-compliance of regulations (such as EU-1169).¹ Complaints can arise for example from people with allergies with regard to inaccurate nutrition information or misleading information on products. The impact of this cannot be underestimated (e.g. damage to brand image and reputation, and non-compliance). Moreover, high data quality increases customer satisfaction by ensuring the completeness and accuracy of product and nutrition information.

The food industry is quickly responding to more demanding consumers and changing habits. They increasingly demand to know what they are actually consuming for health, allergy, economic, environmental, social or even

ethnic purposes. More and more sources are available to customers, making it easier for them to assess the correctness of labels on food nutrition. Incomplete data will be considered as misleading information by customers. For example for jam products the amount of fruits and sugars per 100 grams of the product is not mentioned. This makes it impossible for customers to compare these products with products where this information is stated as required.² The detection of incorrect data will be communicated in a few seconds by (social) media, having an enormous impact on a company's reputation. Along with responding to increasing consumer demands, high quality data also increases supply chain efficiencies through cost and error reduction. This has also been acknowledged by several big players in the Dutch food sector, such as Ahold and Superunie. Suppliers were providing products including data to several retailers, however this information was communicated using different systems, standards and definitions.

¹ <http://eur-lex.europa.eu/legal-content/NL/TXT/PDF/?uri=CELEX:32011R1169&from=FR>

² <http://www.het-nieuweetiketteren.nl/streekproducten-en-consument/>

In addition to increasing customer demands, high quality data is also important for increasing the reliability and efficiency of product transportation and delivery to stores and warehouses

In addition to increasing customer demands, high quality data is also important for the transportation of goods by increasing the reliability and efficiency of product transportation and delivery to stores and warehouses. High data quality of dimension attributes (height, weight) leads to reduced measurement errors of products and can be used to improve efficiency of processes and optimization of transportation through the supply chain. In one example freight trucks were loaded inefficiently because the packaging material of a product was smaller than expected, leading to non-used space in these trucks and therefore higher costs.

The importance of high quality data has also been addressed by government authorities. Mainly health related factors have resulted in stricter legislation and requirements from government bodies. This requires companies to be in control of their data. Specifically the EU-1169 regulation has been introduced to control the importance of data quality in the food sector. More specifically, this EU regulation on the provision of food information brings together different EU rules on general food and nutrition labeling. Regulation is generally intended to protect consumers with several requirements for the food industry sector to provide accurate and complete product and nutrition information on the label. In practice this regulation makes nutrition labeling mandatory, and instructs food manufacturers to provide information on nutrients. Hence, retailers need to have access to the data which is created by suppliers in the supply chain. Obviously retailers have multiple suppliers that use a variety of data transfers standards and want to exchange data at different times. For this reason, a central data source has been set up that functions as a storage facility for both sup-

pliers (deliver and enter data) and retailers (extract data to inform customers with nutrition information). This data pool is governed by a non-profit organization, called GS1.

Over the last couple of years GS1 has made good progress (e.g. data quality programs) but is still facing challenges to achieve the desired level of data quality. Initiatives, lessons learned and activities to improve data quality in the food sector, together with future challenges will be addressed in this article.

Specific challenges of data quality in the food industry

The importance of data quality has been acknowledged across various industries such as financial services, medical devices and consumer goods. More and more industries are subject to increasing customer demands and stricter regulations regarding the use and quality of data. Along with increasing regulation, more and more organizations are transforming into data-driven organizations.

In general, high data quality becomes relatively more important according to the size and complexity of an industry. The supply chain in the food industry is known to be very complex, dynamic and consumer driven. The food industry is also a highly competitive environment. High data quality is therefore a requisite to enable process efficiency and increased customer satisfaction. In practice, this means that companies dealing with product, article or nutrition data are facing the following typical challenges.

Dimensions

Dimension data such as size or weight of materials might be pretty straightforward if you have agreed upon the measurement rules. However, with food products it is not always that clear which dimension is applicable, e.g. is it the length of the product or the width? Try to apply this to a net of oranges to see how difficult this can be. Errors due to incorrect dimension data possibly lead to the inefficient transportation of goods, as trucks and shipping are planned based upon such measures. Consistent definitions of dimensions also contributes to effectiveness of collabora-

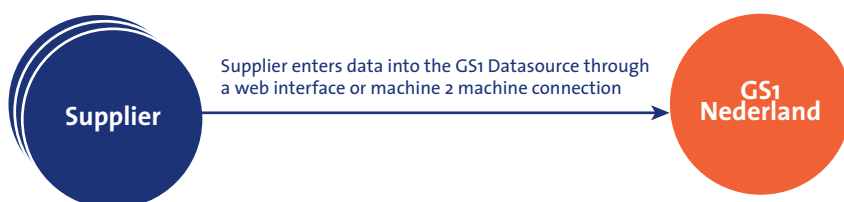


Figure 1. Data entry into GS1 Datasource.

tion between departments and organizations resulting in a more integrated supply chain.

Packaging

Another challenge occurs with the packaging of food products. Do you include the foldable tips of the packaging of a sausage or do you fold it back to have the minimal size? Why bother about this at all and just agree on one way of measuring or the other? Because it actually does have an impact, since the transportation, storage, and locations, are arranged based on this information. Consider the additional costs that a supplier incurs as a result of a half-empty truck as mentioned in the example earlier, caused by calculation errors.

And then we have single articles and bags and boxes and all other kinds of packaging which make things more complicated than they may look at first sight. There may also be several types of packaging for one product, all of which might have different label information and also contain not only text, but also logos which also need to be recorded in text. It can be assumed that logos will be interpreted differently by each person.

Labels

One other typical aspect in the food industry is with regard to label information. Some data only needs to be recorded if it is actually presented on the label, such as the percentage of alcohol. In fact this is a very simple rule, however it is difficult to check or even to automate this. Because while an empty field might indicate a data quality issue, you can only be sure by actually checking the label to see whether this data is or is not available and therefore whether the empty field is valid or not.

A collaborative 'taskforce' approach as a starting point

Some major players in the food industry realized something needed to happen in this area in order to provide

an answer to the challenges and requirements they faced with regards to the quality of product data. In order to get sector-wide support for the data quality 2.0 program a taskforce was appointed. This 2.0 program was introduced soon after the 1.0 program as a renewed and broader approach and because solutions seemed to be necessary. The taskforce is a collaboration of twenty industry leaders consisting of suppliers, retailers and branch organizations. The suppliers are among others represented by Federatie Nederlandse Levensmiddelen Industrie (FNLI),³ Unilever, Pepsico, Proctor and Gamble and Nestle and the retailers by Albert Heijn, Jumbo, Sligro and SuperUnie. Together the taskforce created the outlines, commitment and funding of the data quality 2.0 program.

³ <http://www.fnli.nl/>

At the start of the data quality 2.0 program the taskforce collaborated intensively to create the outlines and principles of the program. There were long debates on topics like the data fields that need to be checked, the involvement of third party data capturing companies and the sample sizes of the physical checks. Now the outlines are clear and the data quality 2.0 program has already started, the taskforce involvement is decreasing, with only a few meetings this year. GS1 is driving and coordinating this program as an independent organization. So who is GS1?

GS1 organization, history and role

GS1 Netherlands is a not for profit company that was founded 40 years ago with the introduction of the barcode.⁴ Albert Heijn himself took the barcode from the US and introduced it to the Netherlands, due to its potential to improve cash register systems in the 70s. In order to fully implement the barcode in the supply chain an independent party was necessary to bring retailers and suppliers together to make the supply chain more efficient. GS1 Netherlands was founded for this purpose.

⁴ <https://www.gs1.nl/over-gs1/over-ons>

Nowadays, GS1 Netherlands offers (international) standards for the unique identification, capture and

In general, high data quality becomes relatively more important according to the size and complexity of an industry

The goal is to improve data quality with a significant impact on the food supply chain and provide more reliable and accurate data to customers

sharing of data. We are still using the barcode, and new central article databases and electronic communications are now available. With the global usage of the barcode GS1 has member organizations all over the world, spread across almost one hundred countries. The global head office in the US is collaborating with companies such as Google, Walmart and Alibaba. Together GS1 works with over a million companies to create industry-wide agreements on identifying, capturing and sharing information.

GS1 Netherlands has grown significantly, GS1 Netherlands now consists of 70 employees, as a result of the data quality program. With a focus area that has changed from sector (projects) managers to an IT and data analytics skilled HR focus.

Data quality program

Over the last couple of years suppliers in the food industry have invested in many good initiatives to deliver product information into the GS1 DataSource. Suppliers aim to provide retailers in the industry with complete and accurate product information. It seems however that despite these good initiatives, data is still in many cases actually delivered incomplete or incorrect. This is applicable for both product/label information and logistics information, such as packaging measurements. GS1 is aiming for 100% accurate and complete product information. In the end, this contributes to the efficiency of processes within the food chain and furthermore provides reliable and accurate information to the consumers.

For these reasons, GS1 has initiated a collaborative data quality program together with CBD (Central Agency of Drugstores),⁵ CBL (Central Agency of Food),⁶ the FNLI (the Dutch Food Industry Federation), some retailers and manufactures. This program is called 'Datakwaliteit 2.0'.

The goal of this program is to improve data quality with a significant impact on the supply chain within the food industry and provide more reliable and accurate data to customers. More specifically, this entails:

- simplifying data entry into GS1 DataSource;
- preventing incomplete or inaccurate data entries;
- correcting incomplete or inaccurate data within GS1 DataSource;

- extending automatic controls within the supply chain;
- making physical controls applicable for both new product information and modified products.

External Data Management Services (hereinafter DMS) play an important role in this solution and will mainly be used for entering or checking the product data in the systems. These DMSs need to be certified and are therefore specialized in order to perform these activities and need to be contracted by suppliers themselves. The next paragraph will detail more the services they can provide with regards to data quality for product data and the associated certificates.

Data management service solution and certificates

The main goal of the project is to guarantee the quality of the product data for which the DMS plays an essential role. They perform data quality controls to assess the quality of the delivered product information by comparing the data in the GS1 DataSource with physical product information. A DMS can also enter logistic and product information into the GS1 DataSource on behalf of the supplier. This product information is excluded from quality controls.

To perform activities related to data quality controls and data entry, a DMS needs to be certified by GS1 and attain a GS1 Quality Mark. Hence, the landscape of the DMS solution can be considered as complex. This has been illustrated in figure 2:

In general, a DMS can attain the following GS1 Quality Marks:

- recording/entry of logistical data and label information (for both the food sector and the health & beauty sector);
- checking logistical data and label information (for both the food sector and the health & beauty sector).

To attain a GS1 Quality Mark, the DMS will be assessed with regard to the following areas:

- *location*: assessment of organizational structures, (data) processes, skills and knowledge of employees;
- *system*: assessment of tooling and systems, more specifically the extent to which the DMS system environment is able to communicate and interface properly with suppliers;

⁵ www.drogistensite.nl

⁶ <http://www.cbl.nl/>

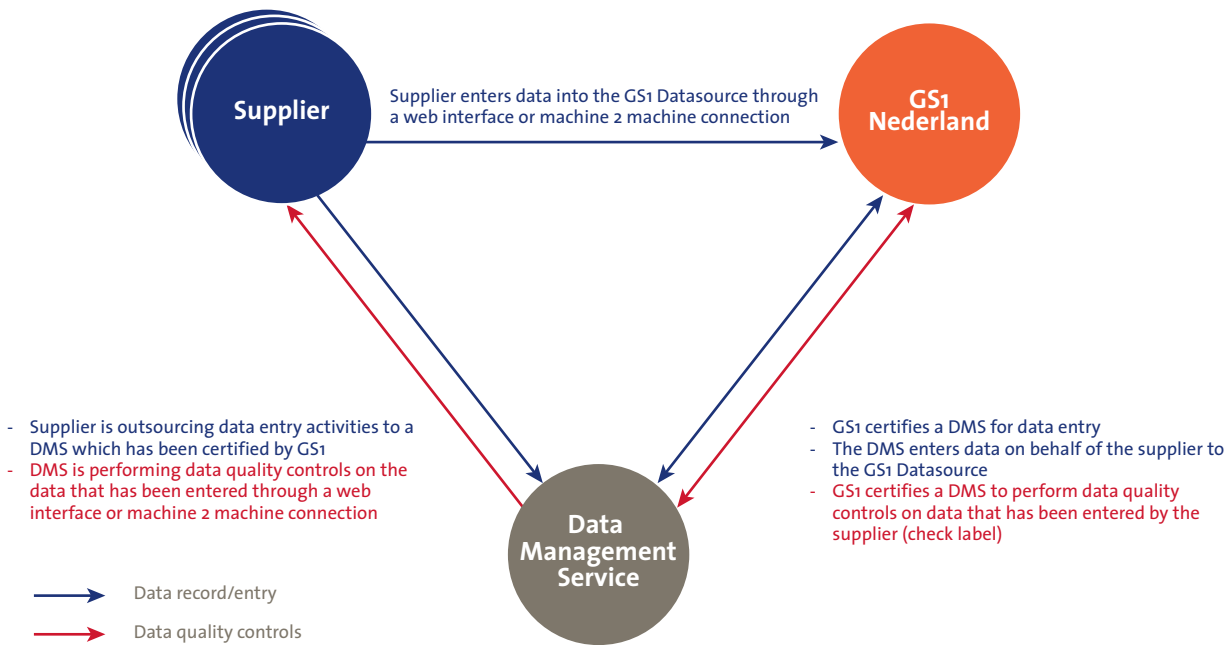


Figure 2. Data Management Solution.

- *quality*: assessment on the output of data entry and data auditing related activities.

The location and system assessment will be performed by an external auditor, while the quality assessment is executed by GS1. Companies need to pass all 3 assessments in order to gain a certificate. Recertification of location and system is on an annual basis, but quality is assessed more frequently depending on the quality of the DMS.

The ultimate goal of the program is obviously to provide high data quality. Hence, a GS1 Quality Mark makes sure that data entry and controls are performed in accordance with the agreements and standards that apply to the Dutch food and drug industry. So for now there is a foundation to reach this goal, but what about the future of GS1?

Future

GS1 is continuously exploring new areas to operate in. As stated at the beginning several triggers did lead to this solution and will provide new triggers when, for example, new legislation and requirements from government bodies are introduced. Also, the scope of this solution, the food and drug industry can easily be copied into other markets which face more or less the same 'challenges' as the health care sector.

GS1 Netherlands is also looking into opportunities to work in a more global way with affiliate organizations in other countries fulfilling the same services. Currently there are already some global meetings and initiatives but a lot is still to be achieved in this matter. As GS1 was and still is in a steep learning curve and continuing to mature, opportuni-

ties are extensively explored into the automation of parts of the solution.

However, the main future impacts will be caused by future developments which cannot be predicted that easily. Who knows when the new 'Albert Heijn' will introduce something totally new and inspiring to this sector requiring services from the GS1 organization? So while GS1 is getting ready for the future, new innovations might require GS1 to innovate as well.

About the authors

A. van Rijswijk is a data quality and analytics manager at GS1 Netherlands. He is passionate about deriving business insights through data analytics. He is currently focussing on data quality of product information in the Food, Health and Beauty sector.

S. Swartjes is a manager at KPMG IT Advisory, Enterprise Data Management. He is an expert in the field of enterprise data management and specifically in master data management.

R.F. van der Ham is a senior consultant at KPMG Enterprise Data Management. He is primarily involved in the financial services sector, focusing on issues around regulatory compliance, data governance and data quality.