



The Global Language of Business

How to improve data quality

Workshop

Data Excellence Day, Cologne

6th of September 2023



Introduction

By Armand Schins

Anti trust

GS1 in Europe will not enter into any discussion, activity or conduct that may infringe, on its part or on the part of its members and participants, any applicable competition laws. By way of example, members and participants shall not discuss, communicate or exchange any commercially sensitive information, including non-public information relating to prices, marketing and advertisement strategy, costs and revenues, trading terms and conditions and conditions with third parties, including purchasing strategy, terms of supply, trade programmes or distribution strategy. **This applies not only to discussion in formal meetings but also to informal discussions before, during or after meetings.**

What we want to achieve

1. Understand importance of data quality & current issues
2. Share insights in current programs & developments
3. Define resolutions & foundation for a EU-way forward



Agenda

1. Introduction
2. Entry-level DQ program – Czech Republic
3. Scaling the manual artwork verification process - GO
4. Comprehensive DQ program – The Netherlands
5. Role of future technologies in DQ
6. House of Commons
7. Closing



Entry-level DQ program

Tomas Tluchor, GS1 Czech Republic

Entry-level DQ program

| | |
|---|---|
| Type of program | Digital – Artwork/image check only |
| Items checked | First three hierarchies of every newly onboarded GDSN supplier |
| DQ Auditor | GS1 Czech Republic |
| Sharing results of DQ inspection | With supplier (data source) only |
| Additional information | Only for GDSN customers of GS1 Czech Republic - every data pool has its own approach to DQ. |



Czech Republic

How does it work – the process



Dimensions – only logical check

- Width = 4 cm
- Depth = 12 cm
- Height = 25 cm



Wrong: Width and Depth
are switched



Correction of issues

- User is informed **by email**
 - Usually text in email body
 - Excel form when many issues are discovered
- Corrections are **checked again**
- **Additional training** offered if needed

Dear user,

Please correct data quality issues below that were found for your items in GDSN.

GTIN 8594182500394 – base unit

- **Tab B.4.: 3517 Product Description** - It should clearly describe the product without abbreviations. It should be brand + sub brand + product type + variant + net content. E.g., "GS1 dark chocolate 85 % 100 g".
- **Tab D.1.: 3733 Net Content** – Value should be 325 ml as on the package.
- **Tab I.1.: 1268 Ingredient Statement** – it should start with the word „Ingredients:“.

Kind regards
GS1 Czech Republic

Pros and Cons

Pros

- Easy to set up
- No extra cost for users
- Making sure new users understand how to work with GDSN
- Improves engagement with users
- Favourable cost/benefit ratio

Cons

- Data quality issues still occur for non-checked products
- Positive effects can decrease over time
- Not addressing user's employees changes
- Slower onboarding process
- Missing technology to improve generation of „Issues reports“ sent to users

Effects of the program

- No comparison „with vs without“.
- Issues found for around 95 % of new users.
- Better than doing nothing.
- More robust DQ program will probably be demanded once GDSN adoption in the Czech market is sufficient.



Scaling up your manual verification process

Nicolas Collignon

Data Quality Programme Manager

GS1 Global Office

Scaling Up The Artwork Verification Process

CHALLENGES

posed by manual verification at scale

- **Increase in the volume** of products and data requires rapid and accurate verification.
- Manual verification becomes **time-consuming and resource-intensive**.
- **Bottlenecks** and delays in the process can hinder efficient data exchange.

STRATEGIES

for efficient scaling

- Prioritize **systematic categorization** of products for streamlined processing.
 - Group products based on common attributes such as category, brand, or complexity.
 - Develop standardized verification templates tailored to each product category.
 - Facilitate efficient allocation of resources and time for focused verification.
- Implement **batch processing** for groups of similar products.
- Utilize trained teams for parallel verification, increasing throughput.

OVERCOMING CHALLENGES

while maintaining accuracy

- **Invest in training** to enhance verification speed without compromising quality.
- Establish **clear guidelines** for consistent and accurate verification.
- Regularly **assess and refine** the scaling strategy based on performance data.

Implementing a Scoring System

Introducing a

SCORING SYSTEM

for brand owners

- Assigning scores based on historical data quality and accuracy.
- Scores influence the level of scrutiny in the verification process.

BENEFITS

of the scoring system

- **Rewards brand owners** with good scores through reduced verification frequency.
- Enhances efficiency by **focusing resources** on brands with lower scores.
- **Motivates** brand owners **to improve** data quality for better scores.

CHALLENGES & CONSIDERATIONS

- Establishing **fair and transparent** scoring criteria.
- **Avoiding biases** that might disadvantage certain brand owners.
- **Balancing the benefits** of reduced verification with the need for accuracy.

Transitioning to Automation and AI

INCREMENTAL SHIFT

from MANUAL to AUTOMATIC verification

- Start by **automating repetitive tasks**, like data extraction from artwork.
- **Gradually** introduce AI for pattern recognition and data alignment.
- **Maintain human oversight** for complex or nuanced verification tasks.

BENEFITS OF AUTOMATION

and AI integration

- **Accelerated verification process** due to enhanced speed and efficiency.
- **Reduction in human errors** through consistent and algorithmic verification.
- **Improved resource allocation** with human reviewers focusing on critical cases.

Gradual

IMPLEMENTATION PATH

- Establish a **roadmap** for gradual automation adoption.
- Invest in AI **training and development** for accurate recognition.
- **Pilot** automated processes on a **subset** of products before full deployment.

CHALLENGES & PITFALLS

to watch out for

- Ensuring AI accuracy and **minimizing false positives/negatives**.
- **Navigating resistance** to technology-driven changes from stakeholders.
- **Maintaining the balance** between automated efficiency and human judgment.

Advanced Automatic Validations and Dashboard Monitoring

Going **BEYOND BASIC VERIFICATION**

- Expand automated processes to **include data consistency** checks.
 - Incorporate **text recognition** for verifying textual product details.
-

AI-POWERED Validations

- AI algorithms analyze data for inconsistencies, outliers, and errors.
 - Alerts are generated for **immediate corrective actions**.
-

DASHBOARD for DQ monitoring

- **Centralized dashboard** displaying verification results.
 - Visual representation of data quality trends and patterns.
 - Facilitates **informed decision-making** and proactive error correction.
-

Empower better **DECISION MAKING**

- **Immediate** identification of data quality issues.
 - **Data-driven insights** for refining verification and data entry processes.
-

Streamlining the Data Quality Journey

- **Recap** of covered topics
 - Scaling up manual verification process.
 - Implementing a scoring system for brand owners.
 - Transitioning to automation and AI.
 - Advancements in automatic validations and dashboard monitoring.
 - Acknowledging the complexities of data quality enhancement.
 - Embracing the potential of technology and innovation to overcome challenges.

Thank you !

If you have any question or wish to discuss these topics further, please reach out:

Nicolas.Collignon@gs1.org

Comprehensive DQ program

Jerry Tracey, GS1 Netherlands

Comprehensive DQ program

| | |
|---|---|
| Type of program | Logical validations (blocking), artwork checks (label) and physical inspections (dimensions). Alternatively it's possible to have data captured externally and to get certified for internal self-checking. |
| Items checked | All new and changed trade items (not pallets) published via GDSN to Dutch food and health & beauty retailers. |
| DQ Auditor | External Data Management Services (checking, capturing) and suppliers (self-checking) certified by GS1 Netherlands. |
| Sharing results of DQ inspection | Detailed reports for suppliers and retailers; status per GTIN as part of GDSN messages. |



The Netherlands

How does it work?

1. Suppliers of Dutch FMCG retailers using GDSN sign up for the DQ program and contract a certified DMS to perform checks or capture data. Alternatively, a supplier can get certified to perform self checks.
2. New/changed GTIN's in GDSN are automatically flagged to be checked (if not captured by DMS or self checked) in MyGS1. Products with an identical package need not all be physically checked.
3. Suppliers send the product or artwork to their DMS.
4. DMS report outcome to supplier and GS1 Netherlands.
5. GS1 Netherlands reports daily status per GTIN to retailers (API), suppliers (MyGS1.nl) and GDSN.



Reasons for choosing this program

- DQ program joint initiative of industry, retail and GS1 Netherlands (2016)
- Driver was the EU 1169 regulation for label information
- Ambition was to include all products and types of information possible
- Retailers demanded that all suppliers would participate in program
- Industry demanded free choice in how (external checking or capturing, self-checking) and with who (variation and competition between DMS)
- Role of GS1 is to facilitate, communicate and support participants



Pros and Cons

Pros

- Almost all FMCG products are in scope (incl. fresh food, pet food, supplements, personal care, cosmetics, detergents, biocides, batteries, etc.)
- Both consumer and case level
- Both new and changed trade items
- Both label information and physical dimensions are checked against product
- Different options and parties to work with for suppliers to participate in program

Cons

- Costs for suppliers and GS1
- Particularly physical checks are a hassle
- Data not always approved in time
- Limited control of GS1 over DMS

Effects of the program

- ✓ >1.410 GLN's in scope -> 1.175 participate
- ✓ 270.000 GTIN's in scope -> 210.000 certified
- ✓ Reliability up from <1% in 2017 to ~80% today
- ✓ First-time-right from under 10% to about 50% today



Additional information

- GS1 Netherlands also offers suppliers and retailers validation reports on product images and allergens. Reports specifically on packaging materials and hazardous materials underway.

Role of future technologies in Data Quality

Jonas Adser, Trade Connectors

Morten Buch, GS1 Denmark



How to improve data quality

Role of future technologies in Data Quality

Jonas Adser
Chief Commercial Officer



jonas.adser@tradeconnectors.org

Role of future technologies in Data Quality - What technologies

Optical Character Recognition

- Converts printed text into machine-readable text that can then be mapped to what format is needed.
- Can automate data extraction from product labels.
- Requires support of a human to validate the result to fully ensure accuracy.
- Potential - Speeds up data entry, reduces errors, and ensures data consistency.



MACHINE LEARNING

Machine Learning

- Creating models which are trained with large datasets to learn patterns and make predictions or decisions based on the dataset.
- Potential is spanning wide from data entry validations, to data suggestions and clean up of obsolete or unnecessary data.
- It is crucial to combine ML with domain knowledge and human expertise to address complex data quality challenges effectively and efficiently.

Role of future technologies in Data Quality - Challenges

The need of receiving the physical product

- In order to be able to capture logistical data the physical product is required which requires more processes.
- Without receiving the physical product data quality checks is limited to label related information.
- Centrally located 3D scanners to capture data, for large producers it might be located at their production site.



Lack of standardization

- Standards to ensure consistency with different data quality checks
- With more automated data quality checks it is an opportunity to share the industry required data quality checks across markets - but this requires standardization.
- Lack of standardization creates a difference across markets - market specific data quality checks that can only be performed and done by a specific entity.

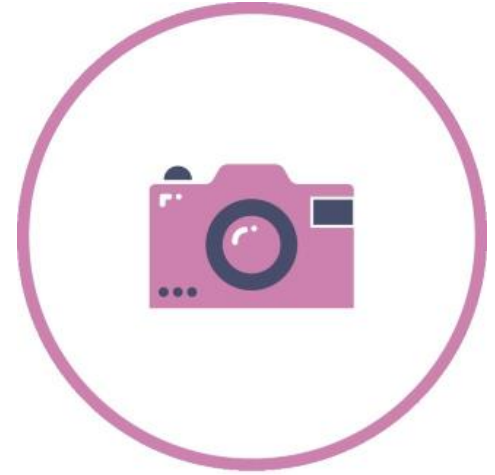
GS1 Denmark



GS1**Trade** Sync



GS1**Trade** Exact



GS1**Trade** Image

2017 – only physical products



2019 – introducing artwork possibility



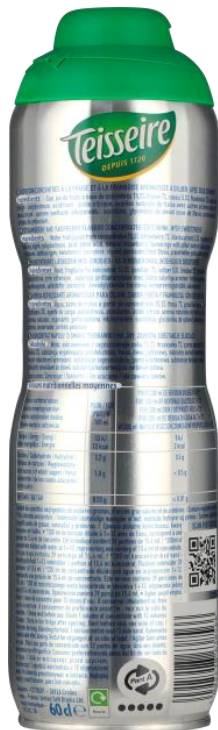
At the moment we need the physical product, when it's introduced.

The aim is to further reduce this need.

Since we have the physical product, we use it to introduce and develop...

Optical Character Recognition

... to help our data quality service.



(EN) Green Tea with honey, with sugars and sweetener - Ingredients: water, sugar, 2% pear juice from concentrate, fructose, honey (0,43%), green tea extract (0,12%), antioxidant: ascorbic acid, citric acid, natural flavouring, sweetener: steviol glycosides from Stevia. Unopened best before end: see bottom of can (UK FBO-WFL: D), The (NL) Groene thee met honing, met suikers en zoetstof - Ingrediënten: water, suiker, 2% perensap uit concentraat, fructose, honing (0,43%), groen thee-extract (0,12%), antioxidant: ascorbinezuur, voedingszuur: citroenzuur, natuurlijk aroma, zoetstof: steviolglycosiden uit stevia. Ongeopend ten minste (SE/NO/DK) Grönt te med honung/honning, innehåller socker och sötningsmedel - Grön te med honning, med sukker og sødestof - socker/sukker, 2% päron-/pærejuice från/fra koncentrat, fruktos/fructose, honung/honning (0,43%), grönt te-extrakt/grön te-extrakt syra/askorbinsyre, syra/syre: citronsyra/citronsyre, naturlig arom/aroma, sötningsmedel/sødestof: steviolglykosider från/fra stevia, öre utgången av/bedst før udgangen af: se botten av burken/se bunnen av boksen/se bunden af dåsen (GR) Πράσινο τσάι με μέλι, με στατικά: νερό, ζάχαρη, 2% χυμός από συμπυκνωμένο χυμό αχλαδιού, φρουκτόζη, μέλι (0,43%), εκχύλισμα πράσινου τσαγιού (0,12%), μέσο όξυνσης: κιτρικό οξύ, φυσικές αρωματικές ύλες, γλυκαντικό: γλυκοζίτες στεβιόλης από στέβια. Ανοίωση κατά προτίμηση πριν από το (HU) Zöld tea és méz ízű szensavmentes üdítőital, cukorral és édesítőszerrel - viz, cukor, 2% körtelék sűrítmenyből, fruktóz, méz (0,43%), zöld tea kivonat (0,12%), antioxidáns: aszkorbinsav, étkezési sav: citromsav, természetes aroma, édesítőszer: szteviol-glikozidok steviából. Minőségét megőrzi: a doboz alján jelzett hónap végéig (LT) Žaliosios arbatos gėrimas su medumi, cukrumi ir saldikliu - Sudedamosios dalys: vanduo, cukrus, 2% kriaušių sultys iš koncentrato, fruktozė, medus (0,43%), žaliosios arbatos ekstraktas (0,12%), antioksidantas: askorbo rūgštis, rūgštis: citrinų rūgštis, natūrali kvapioji medžiaga, saldiklis: steviolio gliukozidai. Geriausia liki (pabaigos): žr. skardinės apacioje (neatidarytos pakuotės) (LV) Zaļā tēja ar medu, ar cukuru un saldinātāju - Sastāvdaļas: ūdens, cukurs, 2% bumbieru sula no koncentrāta, fruktoze, medus (0,43%), zaļās tējas ekstrakts (0,12%), antioksidants: askorbinskābe, skābe: citronskābe, dabīgs aromatizētājs, saldinātājs: steviolglikozīdi. Ieteicams līdz: skat. uz iepakojuma (EE) Roheline tee meega, suhkrate ja magusaineiga - Koostis: vesi, suhkur, 2% pinnimahl kontsentraadist, fruktoos, mesi (0,43%), rohelise tee ekstrakt (0,12%), antioksidant: askorbiinhape, hape: sidrunhape, looduslik lõhna- ja maitseaine, magusaine: stevia stevioolglükosiidid. Avamata parim enne lõppu: vaata purki põhja 990077

DESIGN REGISTERED | SHAKE WELL

500ml e



• NO ARTIFICIAL PRESERVATIVES

AriZona Europe Coöporative U.A. , PO Box 15678, 1001 ND, Amsterdam, the Netherlands | www.drinkarizona.com

| Nutrition Information / Voedingswaarde / Næringsinnhold / Πληροφορίες Διατροφής / Atlagos tápérték tartalom / Maistingumas/ Uzturvērtība / Toitumisalane teave | | per 100 ml: |
|--|---|-------------|
| Energy / Energie / Energi / Ενέργεια / Energia / Energiení verté / Energētiskā vērtība / Energlasisaldus ————— 78kJ / 19 kcal | Carbohydrate / Koolhydraten / Karbohydrat / kuhhydrat / Υδατάνθρακες / Szénhidrát / Angliavandeniai / Ogljhidrati / Süsivesikud ————— 5 g | |
| Fat / Vetten / Fett / Fedt / Ķīmipē / Zsír / Riebalai / Tauki / Rasvad — 0 g | of which sugars / waarvan suikers / hvorav sukkerarter / heraf sukkerarter / εκ των οποίων οι σάκχαρα / amelyből cukrok / iš kurių cukry / tostarp cukuri / millest suhkrid / ————— 4,7 g | |
| of which saturates / waarvan verzadigde vetzuren / hvorav mettet fett / Heraf mættede fedtsyrer / εκ των οποίων κορεσμένα / amelyből telített zsírsavak / iš kurių sočiųjų riebalų rūgščių / ————— 0 g | Protein / Eiwitteen / Protein / Πρωτεΐνες / Fehérje / Baltymai / Olbaltumvielas / Valgud / Proteine ————— 0 g | |
| to starp piesātinātās taukskābes / millest küllastunud rasvhapped ————— 0 g | Salt / Zout / Salt / Alkəm / Só / Druska / Sals / Sool ————— 0 g | |



Morten Buch
GS1 Denmark

mbu@gs1.dk
+21659698

House of Commons

Let's debate!

House of Commons

- How does it work?
 - 1 question/affirmation
 - Every person in the room chooses to stay on the Yay or the Nay side
 - The two groups take a couple of minutes to debate
 - We move on to the next question
- This session will be recorded for note-taking/minutes purposes

1. Compared to other master data exchange solutions you know, GDSN provides the best data quality

Nay



Yay

2. Could GS1 do more to enable organisations achieve higher data quality?
What & how?

Nay



Yay

3. External audits should focus solely on legally required data

Nay



Yay

4. All attributes my company works with are equally important and their quality level must be the same

Nay ← → **Yay**

5. GS1 should focus on certification of data management services, not run them

Nay



Yay

6. Data quality is an important part of product quality control processes in my company

Nay



Yay

7. Data quality programs in different markets should be mutually recognized

Nay



Yay

8. AI will significantly reduce the need for external audits with physical checks

Nay



Yay

9. Product master data quality in my market/company is sufficient

Nay



Yay

Closing

A stylized orange figure with a circular head and a curved body, appearing to be in a celebratory or dancing pose.

Thank you!