Compatibility of EANCOM & GS1 XML

(in particular regarding master data synchronisation)

Issue 1, Approved, August 2011

Document Summary

Document Item	Current Value
Document Title	Compatibility of EANCOM & GS1 XMLMaster Data Synchronisation
Date Last Modified	Aug-2011
Document Issue	Issue 1
Document Status	Approved
Document Description	In this paper the compatibility of the GS1 EANCOM and the GS1 XML message standards is discussed, with a focus on master data synchronization.

Contributors

Name	Organization
Hein Gorter de Vries (editor)	GS1 Netherlands
Véra Feuerstein	Nestlé (Switzerland)
Anders Grangard	GS1 GO
Coen Janssen	GS1 GO
Eric Kauz	GS1 GO
Jean-Luc Leblond	GS1 France
Sean Lockhead	GS1 GO
Staffan Olsson	GS1 Sweden
Pere Rosell	GS1 Spain
Gabriel Sobrino	GS1 GO
Jan Westerkamp	GS1 Netherlands

Log of Changes in Issue 1

Issue No.	Date of Change	Changed By	Summary of Change

Disclaimer

Whilst every effort has been made to ensure that the guidelines to use the GS1 standards contained in the document are correct, GS1 and any other party involved in the creation of the document HEREBY STATE that the document is provided without warranty, either expressed or implied, of accuracy or fitness for purpose, AND HEREBY DISCLAIM any liability, direct or indirect, for damages or loss relating to the use of the document. The document may be modified, subject to developments in technology, changes to the standards, or new legal requirements. Several products and company names mentioned herein may be trademarks and/or registered trademarks of their respective companies.

Table of Contents

1. Introduction 4

2. Summary and recommendations 4

3. Principles of message standards 4

4. Compatibility of data elements and messages 5

5. Compatibility in the exchange of master data 6

5.1. Bilateral exchange 6

5.2. Data synchronisation through GDSN 6

5.3. About messages (CIN and PRICAT) and data elements (GDD and TDED) 8

1. Introduction

In February 2011 the GS1 Architecture Group published the ICOM (Integrated Communications) paper, which provides an integrated view on communicating business data with GS1 standards, including how the original GS1 standards and the EPCglobal standards relate to each other.

Additionally, in this paper the compatibility of the GS1 EANCOM and the GS1 XML message standards is discussed, with a focus on master data synchronization (and hence not on transactional messages).

2. Summary and recommendations

GS1 supports two message standards, EANCOM and GS1.XML. The former is the older one, dating back to early nineties, when there was no widespread internet, while the latter complies with more recent web oriented rules. Within the GS1 community the usage of EANCOM is substantial compared to XML, notably because of "installed base". Because the syntaxes are constructed in different ways and also because the way business processes were moulded thereupon is different, the functionalities offered by two standards are not fully aligned. Such alignment never was stated as a prerequisite.

In this paper it is argued that such compatibility should not be regarded as a goal by itself, but rather must be assessed based on user requirements, when applicable. Always a trade off must be made between the added value of modifying both standards to comply with new requirements versus the development and implementation costs thereof.

In particular, with respect to the synchronisation of ***master data*** through GDSN and connected data pools (which is based upon GS1.XML), a datapool is free to offer non-XML options (such as the use of EANCOM) to their connected companies, as long as the data content and quality comply with the overall GDSN standards. Whether it is warranted to modify both GS1.XML and EANCOM standards to encompass new functional requirements, is to be determined when a request is made; this should not be the general rule. However, especially for the longer term, it is recommended to use GS1.XML.

3. Principles of message standards

For the exchange of business data between trading partners, GS1 standards use two different syntaxes:

- ***EDIFACT*** (Electronic Data Interchange For Administration, Commerce and Transport)). This is managed by UN/CEFACT (the United Nations Centre for Trade Facilitation and Electronic Business).
- ***XML*** (eXtensible Markup Language). This is a set of generic rules for encoding electronic documents, created by W3C (the World Wide Web Consortium), which is widely used in all sorts of applications. UN/CEFACT is basing its UN XML rules on these, as is GS1 with GS1.XML. These two share many rules, but also have specific rules.

The *characteristics* of these syntaxes are:

- Messages contain ***data***, which are stored apart from the messages, to provide data consistency over different messages (e.g. a date is defined identically in an order and in an invoice, both at the data type level and at the semantical level (e.g. the meaning of "requested delivery date" should be the same in all messages)):
  - In ***EDIFACT*** the *data* are defined in the ***TDED*** (Trade Data Element Directory); *logical groupings* thereof (e.g. how to construct an address out of street name, house number, zip code, town, etc) in the ***EDCD*** (EDIFACT Composite Data Element Directory) and ***EDSD*** (EDIFACT Segment Directory).

- For ***GS1 XML*** the *data* are defined in the ***GDD*** (GS1 Global Data Dictionary). There are some components from *other* standards organisations, for example, the ***UN/CCL*** (Core Components Library); this includes ***ABIE***'s (Aggregated Business Information Entities), which describe *logical groupings* of data and the UN/CEFACT Standard Business Document Header, which defines the contents and parameters of the message.

- The ***standard values*** of data elements are defined in ***code lists***.
- The ***procedural rules*** by which the data and segments are combined to form messages, are defined in ***UMM*** (the UN/CEFACT Modelling Methodology); UN/CEFACT has defined global standards:
  - UN/EDIFACT (based upon UN/TDED, EDSD and other (like code lists))
  - UN/CEFACT XML (based upon UN/CCL and other).
- The *business process based rules* determine the ***functionality*** of the messages at the semantical level:
  - For EDIFACT the ***UNSM***'s (United Nations Standard Messages) offer such functional rules for a large number (over 200) of interorganisational processes (both business and government). These messages are composed of segments, each offering a certain "subfunctionality", which are used in several messages.
  - For XML at the UN level several messages have been developed. These are not yet as extensive and mature as the UNSMs (GS1 and others, like UBL, seek to get their semantics incorporated).
- The GS1 message standards relate to these UN standards as follows:
  - EANCOM messages are *subsets* of a selection of UNSM's; they are constructed according to the UN/EDIFACT syntax and make use of the applicable data definitions in the TDED. However, in practice there is not always a full subset compatibility.
  - GS1 XML messages are *not* subsets of UN XML messages; they are constructed according to the XML Design Rules for GS1, which is partly based on the UN/CEFACT XML Naming and Design Rules and make use of the applicable data definitions in the GS1 GDD (Global Data Dictionary), which is partly aligned with the UN/CEFACT CCL.

4. Compatibility of data elements and messages

- The data definitions of similar data elements in the UN/CEFACT CCL (Core Components Library) and the UN/EDIFACT TDED (Trade Data Elements Directory) are not fully identical.
  - The CCL has not been designed as an upgrade of the TDED and hence these are not fully compatible.
  - The same applies to the GS1 GDD versus the EANCOM subset of TDED since GDD only is consistently applied for BMS development. In other words, the GDD content reflects the terms used in GS1 XML messages. GS1 GDD and the EANCOM subset of TDED are only aligned to the degree that the GS1 XML and EANCOM are aligned. Hence GS1 GDD and GS1 TDED are not fully compatible.
  - This also is true between the GS1 GDD and the UN/CEFACT CCL, while on the other hand the CCL also contains data elements from other sectors than those supported by GS1 standards. This also is the case for TDED versus the data elements used in EANCOM.
- In line with this, the GS1 XML CIN (Trade Item and Catalogue Item Notification)message offers more data elements, and hence a broader functionality, than the EANCOM PRICAT message. This lack of compatibility doesn't only apply to master data but also to transactional messages (orders, invoices, etc).

- Is this a real business issue? Probably not: *such compatibility is not a goal by itself, but should always be assessed versus user requirements*. In other words, from a purist point of view, there is an issue, but the price to achieve full compatibility as such is too high. Standardisation is about making differences manageable, to the extent required; it is not about eliminating differences per se. Hence this is not an issue to do something about.
- From the above it follows that *the respective libraries need not be kept aligned as a goal by itself*; upgrades should only be based upon user required modifications.
  - To what extent changes in EANCOM messages and the accompanying subset of the TDED can and should result in change requests for the UN/EDIFACT UNSM's (of which EANCOM is a subset) and the full TDED, should be evaluated at an ad hoc basis. The same applies to the GS1 GDD and the UN/CEFACT CCL.
  - There is no general rule describing which business requirements need to be compatible across libraries. For example, some sector-specific requirements may not be needed in all libraries (depending on their use), while basic data elements may need to be mirrored. However when a new requirement is submitted it is hardly ever immediately evident what is the extent of the compatibility required.
  - Comparing the CCL to the GDD, to determine the applicability of a change request, is more than just an alignment of data requirements. The data itself is not semantically compatible due to differences in methodologies, for example, naming conventions. Compatibility here would be a challenge compared to EANCOM versus EDFIFACT UNSM's.

5. Compatibility in the exchange of master data

5.1. Bilateral exchange

Originally, in the earlier years of exchanging EDI messages using the EANCOM standard, the source of traded products (notably the manufacturer) could send an EANCOM PRICAT message in a bilateral way, i.e. directly to the buyer of the products. The same communication channel would be used as for transmitting transactional messages such as orders, despatch advices, invoices, etc. Often, however, in practice the master data were exchanged on paper, spreadsheets or other means.

At that time datapools did not yet exist. When these became available, companies using EANCOM PRICAT messages could continue to use these to exchange master data with their data pools (provided that the data pool offered this option). In the 2010 survey of eCom usage, 34 countries reported the use of PRICAT. These exchanges can be both between companies bilaterally and between companies and their data pools; this distinction was not examined.

When GS1 XML messages for master data became available, companies could use these for transmitting master data bilaterally as wellwith their data pool; however, if EANCOM was already used there was no business rationale to change to the new syntax.

5.2. Data synchronisation through GDSN

In the mean time a more sophisticated process had become available, which describes the way datapools communicate with each other, including the choreography of how master data are to be transferred to those other datapools from which data recipients retrieve master data required by them.

This process is master data synchronisation using the **GDSN** (Global Data Synchronisation Network), which has become the recommended way of communication, because of its added functionalities, which contribute to a consistent, optimised manner of making master data available to all business partners. Keeping master data synchronised between the source and all the recipients takes more than just transmitting a (paper or electronic) catalogue containing master data.

The **GS1 GDSN** standards offer an ***ongoing process*** of keeping master data about the same articles ***identical*** between two companies ***over time***, notably when parts of those data (i.e. values of attributes describing the item) change due to (minor) modifications in specifications, not warranting the allocation of a new item number. Also, it was recognised that communicating the data is part of a more complex process, since providing correct data in practice is quite a challenge: that is why "***data quality***" has become a major issue. The **GDSN standards** are made up of different components:

1. The **business data elements** (content or actual "business data") which is defined by the ***GDD*** (Global Data Dictionary), and the GS1 XML Catalogue Item Synchronisation suite of standards, including the GDSN Trade Item which is contained in the main communication message between the data pools: the **CIN** (Catalogue Item Notification) message.

   Note: when using EANCOM PRICAT (bilaterally, outside of GDSN, or to up- or download data from a data pool (c.f. section 5.3)), data as defined in the TDED are used, however also GS1 defined data element qualifiers and the FTX (free text) segment is being made to allow the exchange of those data elements which are not defined in the TDED.

2. The **GDSN Validation Rules**. These are intended to reflect business rules and other logical or target market specific rules that are not covered by XML schema validations. However, not all business rules can be validated, since they can be too complex to do so in an automated way. Actors within the GDSN (Data Pools and GR) are required to implement all validation rules intended for their respective role within the GDSN. The rules are syntax neutral. Any company uploading information into a datapool for further communication via other GDSN compliant datapools is supposed to pass the GDSN validation rules, regardless of the format used to upload the information.

3. The **choreography** ("the way master data is communicated via interconnected datapools"). This is exclusive to GDSN and is based on GDD elements. This includes any processes to communicate any updates on attribute values.

   This process is triggered by the GS1 **Global Registry** (GR), a service offered by GS1. The GR registers and stores a subset of item information and also registers and stores subscriptions to items based on criteria that is the same as the information about the item. The GR matches the subscriptions with the items that meet the subscription criteria, and then forwards this subscription information on so that the rest of the synchronisation process can be completed. After the GR has facilitated the original synchronisation and the item data is synchronized, any *update* is sent from the "source data pool" to the "recipient data pool" without any intervention of the GR.

4. The process to ascertain the **quality** (correctness and reliability) of the master data, which is decribed in the **Data Quality Framework** (DQF).

   Note: the DQF is not mandatory when using GDSN; however it must be strongly recommended to be used. Also, the DQF can and should also be applied when exchanging master data bilaterally, irrespective whether EANCOM or GS1 XML is used.

5. A **Certification** process, describing how to check that Data Pools are in compliance with the standards. All data pools must successfully complete this Certification process before being connected to the GS1 Global Registry and the other certified data pools.

6. **Global Product Classification** (GPC), which is a system that gives trading partners a common language for grouping products in the same way, everywhere in the world.

   Note: GPC is meant to be used within all applicable components of the GS1 System including GDSN, GS1 XML and EANCOM.

5.3. About messages (CIN and PRICAT) and data elements (GDD and TDED)

In both EANCOM and GS1.XML messages are available regarding master data:

1. The **CIN** (Catalogue Item Notification) message is part of the GS1 XML messages. This is based upon the **GDD** (Global Data Dictionary).

   The CIN is used **between Data Pools**. GDSN Standards are *not* normative regarding the communication between the source (= supplier) of the master data and its Data Pool, nor between companies bilaterally (P2P). This is in line with the original vision of GCI (the Global Commerce Initiative, a precursor to the Consumer Goods Forum), where the master data communication between a company and the Data Pool was outside of the standard. However, in practice there are data sources and data recipients exchanging information with the data pools using the GS1 XML CIN message.

   The CIN message is constructed by using a series of **Trade Item** messages, built to represent a **logistical hierarchy** (for example, consumer unit that are packaged together and, in turn, those packages are put together in a case. Each of these 'levels' would have its own item number and its own set of information about it. It is mandatory to use the CIN message in between GDSN data pools. However, companies can use the CIN message with their data pool (and also bilaterally) and also the Trade Item message.

2. The **PRICAT** (Price Catalogue) and **PRODAT** (Product Data) messages, plus the **PROINQ** (Product Inquiry) message, are part of the GS1 EANCOM standard. They are based upon the EDIFACT EANCOM **TDED** (Trade Data Elements Directory). They are designed for **bilateral**(P2P) exchange; according to the EANCOM specifications they may also be used to upload or download data with a data pool (some pools offer this option, which is used in practice).

It should be noted that, between a user company and a data pool ***any format*** may be used, depending on the offering of the data pool provider: CIN& TradeItem and PRICAT (described above) and also PRODAT, spreadsheet or web-interface (notably for SME's which don't have systems capable of producing an electronic master data message).

It is recommended to use the TradeItem or CIN messages when communicating with a data pool, while offering the freedom to use other options (i.e. *it is not to be discouraged* to use these other options to communicate between a user company and its data pool (i.e. outside the GDSN network)).

To take optimal advantage of automated business systems it is best to use GS1 XML or EANCOM for all data exchanges. However, companies not capable of structured EDI, or in early stages of implementation of GDSN, may use web-based interfaces and perhaps even spread sheet files (such as Excel). The latter, however, is not a viable long term solution, while a web interface requires manual input of data, which is error prone and does not make use of automated systems, if available. That's why it is recommended that, ideally, companies best use the TradeItem or CIN messages to communicate with their data pool.