

Working Paper No. 3

Topic (i) XBRL and business reporting

**UNITED NATIONS
STATISTICAL COMMISSION and
ECONOMIC COMMISSION FOR EUROPE**

CONFERENCE OF EUROPEAN STATISTICIANS

**EUROPEAN COMMISSION
STATISTICAL OFFICE OF THE
EUROPEAN COMMUNITIES
(EUROSTAT)**

Work Session on Electronic Raw Data Reporting
(Geneva, 6-8 November 2006)

THE DUTCH TAXONOMY PROJECT AND XBRL Electronic Raw Data Reporting and XBRL

Johan Lammers, Statistics Netherlands



Statistics Netherlands

Division of Methodology and Quality
Methodology Heerlen

THE DUTCH TAXONOMY PROJECT AND XBRL

Johan Lammers

Summary: For efficient exchange of data, standardisation is of paramount importance. Efficient exchange of data requires semantic standardisation at a national and international level. Standardisation of both financial data and processes of exchange of data between businesses and government can increase the quality of reporting, reduce the time needed and reduce costs for multiple stakeholders. Statistics Netherlands is one of the stakeholders and plays an active role in the Dutch Taxonomy Project. Knowledge of the data to be exchanged, of the processes that accompany the exchange and the use of data and of the methods to use within these processes, help to play this role. This snapshot describes the state of play.

Keywords: Dutch Taxonomy, Response burden, XBRL

Project number:

BPA number:

Date:

12 October 2006

1. Introduction¹

For efficient exchange of (financial) data, standardisation is of paramount importance. Efficient exchange of (financial) data requires semantic standardisation at a national and international level. Standardisation of both financial data and processes of exchange of data between businesses and government can increase the quality of reporting, reduce the time needed and reduce costs for multiple stakeholders.

A shift of paradigm from government to private sector can strongly benefit successful implementation of semantic and process standards. Useful standards are linked closely to the existing processes and administration of businesses. Standardisation can enable companies to easily report to multiple organisations, re-using information derived from one administration.

In the Netherlands, a Dutch XBRL-taxonomy (eXtensible Business Reporting Language) will be implemented in the financial reporting chain, thus achieving cross domain standardisation in the field of annual accounting, taxation and economic statistics. The main driver is reducing administrative burden (the costs of obligatory retrieving and delivering information to government) for companies. A strong political support is not enough to make this project successful. Each member in the financial reporting chain needs to acknowledge and experience the advantages of working with the taxonomy. The Dutch Taxonomy Project stands for timely delivery of the taxonomy and moderates the interests of the members in the chain in both applying the taxonomy and creating the necessary process infrastructure for exchange of information between government and companies.

One of the members in the chain is Statistics Netherlands. The standardisation of financial data and of the processes to exchange them demands for adaptation. XBRL has a major influence on these changes. Therefore, the R&D-program of Statistics Netherlands studies several opportunities to use XBRL within the processes to collect and to validate data, with the aim to implement the methods from 2006 onwards.

2. The Dutch Taxonomy Project

The Ministry of Justice and the Ministry of Finance in a joint effort intend to reduce the administrative burden of businesses by using the open standard XBRL, to simplify the composing of financial reports and exchange of financial data. The

¹ General introduction of this topic is based upon Harm Jan van Burg and Mariette Lok in 'Advantages of cross domain standardisation of financial data', September 2006.

activities should result in a yearly reduction of €350 million for businesses in the Netherlands as from 2007 and of €420 million from 2008.

An important aspect of the project is creating the right environment for this type of innovation, and thereby for a large scale use of the taxonomy. This includes informing and consulting all relevant stakeholders, among which companies, accountants, auditors, intermediaries (and their associations), software suppliers, governmental bodies and ministries. Partnerships are formed between private businesses, for instance software suppliers and intermediaries (private – private). Other partnerships include intensive cooperation between different governmental bodies (public – public). The project is transparent and open to all potential participants. Review of the taxonomy takes place in national and international bodies.

The Dutch Taxonomy Project is responsible for the architecture of the taxonomy and moderates the process of building it. The governmental bodies themselves are responsible for the domain taxonomies. Professionals are engaged to perform the necessary work.

Another important aspect is a constant drive to accomplish an authorized and robust taxonomy and a necessary process infrastructure in time and within budget. By delivering the products in time and within budget the Dutch Taxonomy Project shows the stakeholders that reducing the administrative burden by optimizing the reporting chain is an unstoppable process. The stakeholders are triggered to change their business models, because they get the assurance that the preconditions will not be changed over time and therefore the investments are preserved.

The scope of the Dutch taxonomy includes taxes, annual accounts and statistics. Thus a cross domain standardisation in the area of financial reporting is achieved. At this stage the taxonomy is tested by governmental agencies, auditors and software suppliers. Broad scale use of the taxonomy is foreseen in 2007. The use of the Dutch taxonomy will be voluntary. Potential future additions to the taxonomy could include data elements in the field of education, health, banking and agriculture.

3. Elaboration of the taxonomy

In order to create a useful version of the Dutch Taxonomy, the project was carried out in several phases. The first phase concentrated on listing the concepts as far as they are relevant for the regulators and on the architecture of the taxonomies. After that scope was reallocated to the harmonisation of the concepts and on the standardisation of the context. This resulted in the creation and publication of version 0.1 of the Dutch Taxonomy. This version was used to work out several proofs of concept by a large variety of contributors. The taxonomies were iterated based upon the experiences from the proofs of concept and the project expanded the focus to the whole chain of financial reporting. Special attention was given to the standardisation of the process and to creation of the conditions for the

implementation of the services. Eventually accountants, software developers, regulators and enterprises were asked to commit to the concept and to sign a covenant that has to lead to a fruitful introduction of the Dutch Taxonomy from January 1st 2007.

3.1 Domain taxonomies

As a first step in the creation of the Dutch Taxonomy, all three regulators compiled their own domain taxonomy. In these taxonomies focus and effort was placed at the construction of a list of concepts. For the statistical domain the scope was concentrated on financial variables for four types of statistics: short-term (monthly or quarterly) statistics in line with the STS-regulation, annual statistics for industrial production (SBS-regulation), annual financial statistics about enterprise groups in the non-financial sector and investment statistics.

The restriction to financial variables resulted in the situation that only a part of the concepts for our annual industrial production statistics was included in the domain taxonomy. The remaining part has to be dealt with in the future.

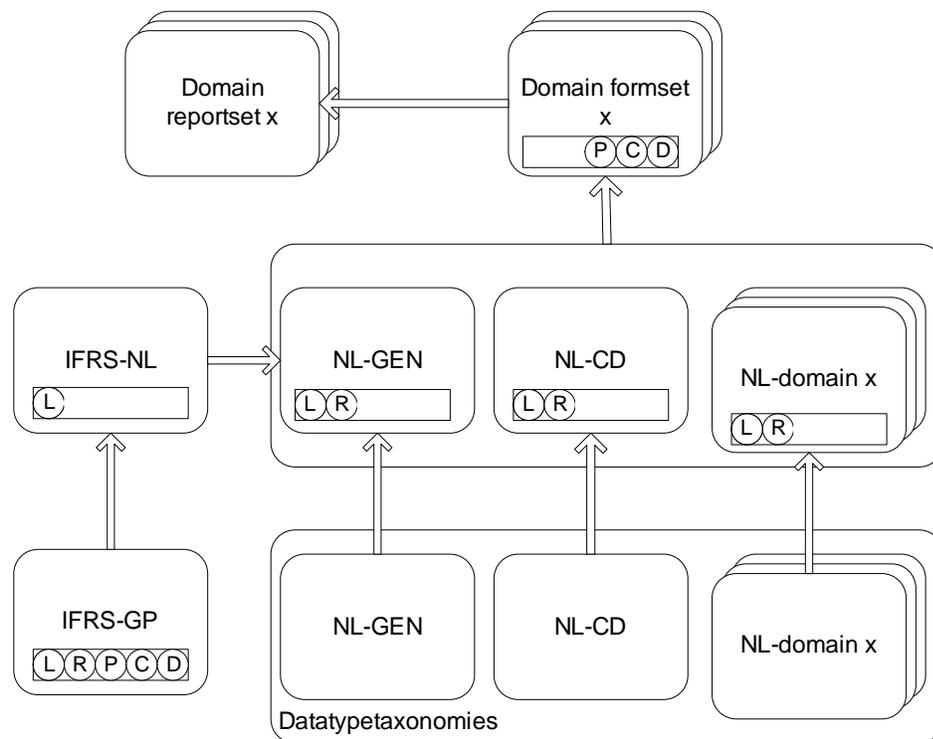
Creation of the domain taxonomies implicated standardisation and normalisation of the concepts. Only a limited internal harmonisation and a limited comparison with IFRS were established.

3.2 Architecture of taxonomies

The architecture contains the design of all relevant parts of the taxonomies and their relations. Based upon the domain datatypes and the domain concepts, a generic taxonomy was created for generic datatypes and another for generic concepts. The NL-GEN was related to the IFRS.

Besides the concepts for the financial variables, a taxonomy was created for contextual information, the NL-CD (Common Document).

To get a better alignment with the contemporary ways to interchange data, report sets were introduced. A report set is a domain-specific view on all the concepts. For the statistical domain, a report set contains all the concepts for a specific questionnaire. This report set uses one or more form sets; blocks of questions. The form sets are domain-specific too and can be used by multiple report sets. Within each taxonomy the link bases have to be determined. The next scheme indicates which taxonomy has a link base for labels (L), references (R), presentations (P), calculations (C) and definitions (D). The arrows in the scheme indicate imported taxonomies.



3.3 Standardisation of context

Until this point, only the content of the message is defined, regardless of its context. Defined concepts like ‘turnover’ can be instantiated for several periods, entities, scenario’s etc. Besides the content of an XBRL-message, the context was worked out within the standard. Eventually the result was defined as concepts within the common document (CD) – taxonomies. Five main aspects concerning the context were explored:

- The definition of the entity. All the regulators deal with a peculiar way to define the entity that should be reported on. From the one hand the statistical offices deals with statistical units (or data collection units that are derived from them) and the tax-office deals with tax-units, whose structure is strongly determined by tax-laws and by the preferences of both the enterprise and the tax-office. Finally the entity for the Chambers of Commerce is determined by the fourth EU-regulation concerning annual reports of enterprises. This part of the context is difficult to harmonise in the short span of time to develop the Dutch Taxonomy. More hope for a successful development in this direction is based upon the creation of a central Dutch registration for objects that are used by several regulators. The creation of a central business register is one of them and a version with the necessary quality and authority is expected to be implemented in 2009.
- Periods, duration and timestamps. The main difficulty in this area was the definition of the financial year. Although these elements were not harmonised, the differences seem to be processable.

- Scenario. In this category the state of the report has to be set. This state can be: preliminary, definitive and corrected. The statistical reports can only be definitive.
- Responsibility, signature and contact information. This group of elements is heterogeneous among the different regulators.
- Audit certificate. The Chambers of Commerce use this information. The corresponding concepts were defined by NIVRA, a national institute for accounting auditors.
- Segmentation. For the compilation of regional or otherwise segmented statistics, the information of the individual segment is needed. Main segmentation is based upon sub organisations, regions, products and activities. Although this aspect is important for statistical purposes, the implementation is postponed. It will be possible to implement this part with the introduction of dimensions in the taxonomy.

3.4 Harmonisation

The most challenging part of the Dutch Taxonomy Project is the harmonisation of the concepts. The harmonised, common, concepts appear in the NL-GEN-base. At the end of 2005 and during 2006 the concepts from the domain of the tax-office were compared with the concepts from the statistical domain.

Two problems arose in this phase:

- The conceptual integrity of the concepts. Not all the concepts to be harmonised were comparable due to a lack of integrity. Some concepts from the tax-domain were defined both as balanced and as unbalanced concepts. The exact definition was not relevant as long as the respondent was consistent within the filling of the instances. From the statistical domain the concepts had to be defined more strictly. Statistics Netherlands applied for the definition of an ‘atomic’ structure of concepts. Finally a compromise was established with the intention to create this structure and to effect this situation during the next two years.
- Interpretation of differences between concepts. For this problem the main strategy was to wipe out small differences and to accept all other differences at this moment.

The result of this activity was a NL-GEN-base with approximately 160 concepts, defined according to the atomic structure. And the statistical domain contains about 90 concepts.

3.5 Standardisation of processes

In order to be able to use the standardised, normalised and harmonised concepts, an adequate process infrastructure is necessary. Therefore the Dutch Taxonomy Project is extended with the (sub-)project to create this infrastructure. This part of the project offers the opportunity to the market to implement functionality that supports

the interchanges of the standardised information, including information on the status of the exchange, based on requirements determined beforehand. This is called the service concept. To facilitate these services, the government has to set up an environment that guarantees the necessary conditions: data standards, process standards, access to registers and, if needed, temporary data storage and basic infrastructure.

The organisation responsible for this part of the process is a government-wide co-operation called GBO (Dutch abbreviation for Common Organisation to manage the data collection). One of the systems used by GBO is OTP (a Dutch abbreviation for Governmental Transaction Port)

In 2006 GBO is working on the implementation of a minimal set of functionality. This includes services for authentication, verification and delivery of information to and from enterprises (or intermediaries) and regulators.

The process stated above is to be classified as an external process. For the individual regulators, like Statistics Netherlands, the design and implementation of internal processes is also essential. As well as the migration process to achieve the required, or even the desired situation.

For the internal processes the most important changes concerns:

- Change management towards the taxonomy. For the initial creation of the taxonomy Statistics Netherlands has to be able to compare a version of the taxonomy with the internal metadata system. After this comparison it has to be possible to generate and evaluate change-issues. This part of the process is relying, at Statistics Netherlands, a lot on individual competences. Our metadata system is not yet equipped to fit with the components needed to establish the desired taxonomy. Main differences are situated in the information to fill-in the link roles.
- Prepare data needed by the external processes to collect instances. Information to pre-populate the instances and to accompany the data collection process can be send to GBO. For the first year Statistics Netherlands decided not to use this opportunity. So this process-step has less priority.
- Receive collected instances. It has to be possible to receive the collected instances technically and to interpret the envelope of the information. The processing has to be secure, reliable and performing. Concerning the performance a risk appears in the estimation of the expected amount of instances for the near future.
- Visualise instances (or 'rendering'). The contents of the instances should be visualised. This offers means to interact with the received information. This topic closely relates to electronic forms.
- Transform the instances to the data structures of our internal dataprocessingsystems. The structure of an instance depends on the definition of the taxonomy and on the specifications of XBRL. These specifications are based

on standardisation of the exchange of reports. The internal processing systems are designed for different purposes and therefore using different data models.

- Incorporate the stream in our existing Customer Relation Management System. Within our CRM system two steps are relevant for the introduction of XBRL as a separate mode to respond: the announcement of the survey and the handling of the XBRL-response. The announcement has to contain an invitation and an entrance to use XBRL. Adequate handling of the response is essential for the follow-up of the process in case of non-response (sending reminders, enforcement of legal obligations etc)
- Validate instances. The quality of the instances is one of the key factors for the efficiency and effectiveness of the data collection process. Validation rules are one of the means to express and measure the desired quality. Incorporating the validation within the process can be realised in several steps. Generally spoken, validation should take place as early as possible. This conflicts with the availability of referencedata to execute validation rules. As the process progresses, more referencedata will be available. Another aspect is the standardisation of validation rules. The more these rules can be expressed in common standards (like XBRL or XML), the more opportunities arise to incorporate the rules within various steps of the process. Finally, the implementation of validation rules should be balanced with regards to performance issues and with the risks to use rules that are more strict then advisable.
- Spontaneous response handling. The use of XBRL can result in spontaneous messages from enterprises. This response is not expected when the enterprise was not invited to respond. Spontaneous response deviated from traditional response-handling, it has its own requirements for storage, methodology and processing.
- Mixed sourcing. The scope of the Dutch Taxonomy is limited to a part of the concepts we need for our statistics. In order to instantiate a complete set of concepts, sources have to be combined.

Besides these processes, Statistics Netherlands has offered to execute the maintenance of the Dutch taxonomy. This role arose out of the function that Statistics Netherlands serves as a national bureau of standards. GBO is responsible for this (common) process, but the execution of this process is still under negotiation.

Maintaining the taxonomy implicates several additional steps concerning the handling of change-issues, taxonomy-validation and taxonomy-versioning.

4. XBRL research

The R&D-program aims to research and implement methods on behalf of the improvement of the statistical process or the statistical product. XBRL is a set of rules with the aim to standardize the interchange of financial business reports. These rules originate from three sources: information technology, information modelling and accounting. From a methodological point of view, the information modelling- and accounting aspects are important. Methods that will be used by Statistics Netherlands should be based upon open and widely accepted standards as much as possible. This is one of the main reasons to research methods related to the use of XBRL within our statistical process in general and within the datacollection process specifically. The work is directly related to the Enterprise Architecture of Statistics Netherlands.

4.1 Validation²

The work recorded in this project constitutes a sound basis for the design of an XBRL workflow system. It is important to realise that the results generalised here were realised by studying one statistical questionnaire and taxonomy (SBS-taxonomy for wholesale-activities). It is clear how the various kinds of checks can be executed on an XBRL instance. A large part of the concept-related checks can be recorded within the XBRL specifications and the formula specifications. For context-related checks and concept-existence checks non-XBRL techniques, such as XSLT, must be used. An XBRL workflow system must therefore at least offer the facilities to apply both solutions.

The only sort of checks for which it seems necessary to use a specific software program are the rules that check for the presence of possible ‘multiple-of-thousand’ errors. Further studies will show whether it is possible to carry out such checks during the creation or analysis of XBRL instances in another way. Possible solutions are: checks during data entry, a complex XSLT or a passage in the message implementation guide

As the research took place in 2005, it made use of version 0.1 of the Dutch Taxonomy. Besides the results on the application of validation rules, a major feedback could be given to de Dutch Taxonomy Project as well as to the specifications of XBRL 2.1.

Further results were a classification of validation rules and a huge jump in our knowledge to generate instances and test cases.

² Daas, P., Stroom, A. (2005), ‘XBRL 2005: Taxonomy and instance creation and control’, <http://www.cbs.nl/nl-NL/menu/methoden/research/discussionpapers/2005/2006-01-x10-pub.htm>

4.2 Data transformation

This research studied the options to transform XBRL-instances to various data models that exist at Statistics Netherlands.

For this aim a proof of concept was worked out for the statistics within the scope of the Dutch Taxonomy Project.

The data of valid XBRL-instances can be stored in the database structures Statistics Netherlands uses in SQL-server, Access and Blaise. The storage is independent of the taxonomy. No data will be lost and all the relational information is transformed accurately.

This result also has a promising side effect: the development of functionality to import XBRL-instances in existing electronic forms. These forms use the Blaise-data model. Importing available data from existing instances support the re-use of these data and relieve the (negative) perceptions about the response burden caused by statistics.

4.3 Generation of electronic forms

Based on a valid XBRL-taxonomy we plan to research in the fourth quarter of 2006 whether it is possible to generate an electronic form. This form should apply to our requirements about the design of electronic forms.

We will base this research on the report set for our annual statistic on industrial production.

The taxonomy for the annual statistic on industrial production will contain some 75 concepts of which roughly 60 will have to be filled in. On the basis of a workable version of the taxonomy for these statistics (which is part of the Dutch Taxonomy) a proof of concept has to be worked out.

In scope of the required proof of concept are a few elements:

1. Template development
2. Forms generation
3. Pre-population
4. Form distribution
5. XBRL import
6. Submitting

The form generation process should ideally be one that can be regenerated - thus - if the XBRL is adjusted, then the form needs to be adjusted as well automatically.

4.4 Relationship with Metadata

One of the crucial aspects on working with taxonomies is the relationship with internal metadata storage. As the taxonomy expresses the metadata of the information Statistics Netherlands wants to interchange with their respondents, it is

important that the corresponding concepts, types, attributes and link roles are consistent with the internal metadata. This relationship works in two ways: is it possible to generate a valid taxonomy using our internal metadata model and is it possible to store the data from a valid taxonomy in our internal metadata model?

Of course, it would be preferable to convert from one to the other metadata model automatically.

This project is planned for 2007.

5. Finally

Developments in XBRL and in the creation of a taxonomy like the Dutch Taxonomy are evolving rapidly. This document offers a snapshot. A glimpse at the realisations, at the state of affairs and at the plans for the future.

XBRL promises to play a large role in the exchange of information. Efficiency and effectiveness can not rely on the introduction of a new method or a new technique by itself. This development has to be a part of the business. Looking at the information supply chain, it even has to be a part of the business of all the participants.

Co-operation is opportune to improve the chain. The Dutch Taxonomy Project is carried out to realise this co-operation and to harvest. XBRL only helps.