Data collection strategy

XBRL

Piet Daas and Marko Roos

Statistical Methods (201105)

Statistics Netherlands

The Hague/Heerlen

Explanation of symbols

.	= data not available
*	= provisional figure
**	= revised provisional figure
x	= publication prohibited (confidential figure)
–	= nil or less than half of unit concerned
–	= (between two figures) inclusive
0 (0,0)	= less than half of unit concerned
blank	= not applicable
2010–2011	= 2010 to 2011 inclusive
2010/2011	= average of 2010 up to and including 2011
2010/'11	= crop year, financial year, school year etc. beginning in 2010 and ending in 2011
2008/'09–2010/'11	= crop year, financial year, etc. 2008/'09 to 2010/'11 inclusive

Due to rounding, some totals may not correspond with the sum of the separate figures.

Publisher
Statistics Netherlands
Henri Faasdreef 312
2492 JP The Hague

Prepress
Statistics Netherlands - Grafimedia

Cover
TelDesign, Rotterdam

Information
Telephone +31 88 570 70 70
Telefax +31 70 337 59 94
Via contact form: www.cbs.nl/information

Where to order
E-mail: verkoop@cbs.nl
Telefax +31 45 570 62 68

Internet
www.cbs.nl

60165201105 X-37

Table of contents

1. Introduction to the subtheme

The eXtensible Business Reporting Language (XBRL) is a reporting standard used by Statistics Netherlands to collect data from government bodies and businesses. This document describes how Statistics Netherlands uses XBRL.

1.1 General description and reading guide

1.1.1 Description of the subtheme

XBRL is an open standard for the compilation and electronic communication of business and other data (XBRL, 2006a; Hoffman, 2006). XBRL is an XML-based language (XML stands for eXtensible Markup Language). XBRL supports the production of reports based on a shared dictionary of terms and their interrelationships. The terms and relationships are defined in a taxonomy. XBRL messages (which contain the data) are referred to as instances. The XBRL standard is an excellent basis for communicating business and other data between, and for compiling reports from, diverse systems (Lammers, 2006). The above indicates some of the reasons why XBRL is of great importance to the government as a communication standard for reporting business data (NTP, 2007a). Depending on the effective integration in the accounting software used by business, XBRL message compilation can be completely automatic, making XBRL an important factor in reducing the administrative burden in the Netherlands.

1.1.1.1 Origin

XBRL was developed in 1998 by Charles Hoffman and the American Institute of Certified Public Accountants (Hoffman, 2006; XBRL, 2007b). They created a list of benefits of using an XML-based financial reporting standard. According to their opinion, XBRL is:

- a global standard, independent of software or platform;
- a way to improve data quality because it eliminates redundant processes, such as data re-entry;
- simple exchange of data between computers, and analysis, possible additional processing, selection and reporting of the data, based on a single taxonomy (i.e. a definition of elements, structure, interrelationships and content of the data elements to be reported);
- an opportunity for substantial cost savings because data in taxonomy-based XBRL format can be selected depending on need or demand.

1.1.1.2 XBRL in the Netherlands

XBRL's benefits were reason for the Dutch government, led by the Ministry of Justice and the Ministry of Finance, to adopt XBRL as a way of reducing the administrative burden, for businesses in particular (Lammers, 2006). The

government's aim with XBRL and the Dutch Taxonomy (NTP, 2007a) is to ease and reduce the expense of compiling and communicating financial data by standardizing data exchange in combination with information and communication technology (ICT). Statistics Netherlands has formalized its adoption of the approach by signing a covenant of cooperation (Lammers, 2006; Roos, 2006).

1.1.2 Guide to the reader

Since XBRL is relatively new, this document starts with a brief discussion of the technique and defines the terms used. The various XBRL files and how they are used are then discussed in greater detail. The information presented provides the background for understanding the methods described below. XBRL is currently still in development, and this document accordingly suggests the directions in which the standard is likely to be extended. It is very likely that Statistics Netherlands will be using at least some of these new developments.

1.2 Scope and relationship with other themes

This document describes the use of XBRL by Statistics Netherlands. Despite a deliberate restriction of the level of detail of XBRL technical aspects, this document provides a necessary minimum explanation of the various XBRL files and their use for a satisfactory comprehension of the subject.

1.3 Place in the statistical process

XBRL is used for both storing and documenting data, and is therefore part of the input process: the data collection part. The end point of the use of XBRL in the input phase is the conversion of XBRL data to a different data format, or the storage of the data in a database. The XBRL data will subsequently enter the data processing operations of Statistics Netherlands. Besides data, XBRL also defines the associated metadata, which is a special feature.

1.4 Definitions

Term	Definition
XBRL	eXtensible Business Reporting Language. An open XML-based standard for exchanging financial data.
Attribute	A descriptive element that adds information. Used, among other things, with concepts.
Balance type	Indicates whether a given item is a debit or credit amount.
Concept	A defined variable in an XBRL taxonomy. There are two types of concept: items and tuples.
Discoverable Taxonomy Set (DTS)	A set of cross-referenced taxonomies.
Fact	A fact expresses the value of an item in an instance.
Formset	A layer in the Dutch Taxonomy in which concepts are structured into internally consistent blocks (forms).
FRIS	Financial Reporting Instance Standards. Additional requirements for an XBRL instance.
FRTA	Financial Reporting Taxonomy Architecture. Additional requirements for an XBRL taxonomy.
IFRS	International Financial Reporting Standards. A set of

	agreements about how companies present their business data.
Instance	XBRL message with data. Always based on a taxonomy or DTS.
Item	A concept that may contain data.
Linkbase	A file containing the relationships between individual 'concepts' and/or documentation and/or labels. A linkbase is part of a taxonomy.
Namespace	A namespace denotes the unique identifier of a document. Is also used to identify the vocabulary of a document.
Reportset	Reporting layer in the Dutch Taxonomy. The reports have one or more formsets.
Schema	File with the metadata definition of XBRL concepts. Is part of a taxonomy.
Taxonomy	Files recording the XBRL metadata definition and the interrelationships between the metadata.
Tuple	A concept that serves as a container for other concepts (which may be items or tuples).
Validation	The process of checking an XBRL document for compliance to the specifications.
XML	eXtensible Markup Language. A standard for defining formal markup languages that represent structured data as flat text.
XSLT	eXtensible Stylesheet Language Transformation. An XML-based language for translating XML-type files into other formats, or differently structured XML documents.

1.5 General notation

The basis of XBRL is the 'taxonomy', which contains the definitions of the terms to be reported. A typical property of an XBRL taxonomy is the extremely rich set of attributes of the variables defined (which are referred to in XBRL as concepts). An XBRL message is called an instance which is always based on a taxonomy. An instance contains the data to be reported.

1.5.1 Taxonomy

A taxonomy consists of an XML schema and files that provide the relationships between the concepts in the schema (XBRL, 2006a). A schema defines the variables to be used in the taxonomy; variables are referred to in XBRL as concepts. There are two types of concept; items and tuples. Unlike tuples, items may contain data. Tuples may contain items or other tuples, but cannot themselves contain data.

A tuple is a container for other concepts (which may be items or tuples). An example of a tuple is an address. The tuple 'address' serves as a container for the following items: street name, house number, house number qualifier, postcode and town. Defining the tuple 'address' in a taxonomy guarantees that a report will present all the necessary components of this concept (street name, etc.) to the user.

Besides a name, items have several attributes, such as data type, period and balance type, which are also specified in the schema. Some frequently occurring data types are 'string' (for text), 'date' and 'monetary' (for financial data). However, it is also possible to define custom data types, e.g. to create menus or to define input constraints. The period of an XBRL concept indicates whether it is a stock variable or a flow variable. A stock variable has a value that is measured at one specific time

(e.g. ‘value of inventories’) whereas a flow variable has a value measured over an interval of time (e.g. ‘revenue’). A concept’s balance type indicates whether it is a debit or credit amount. However, balance type has to be defined only for concepts that relate to a company’s statement of financial position (balance sheet).

Files that contain the relationships between concepts are known as linkbases. These relationships are referred to as hierarchies. Linkbases contain information on the interrelationships between concepts and the relationships between concepts and documentation. Examples of interrelationships between concepts include the definition, calculation and presentation linkbases.

- The definition linkbase describes the logical relationships between concepts, such as their mutual dependence.
- The calculation linkbase describes simple arithmetic relationships between concepts, e.g. which concepts are included in the value of the total concept.
- The presentation linkbase records the presentation structure of the concepts.

The relationship between concepts and documentation is given in the label and reference linkbases.

- The label linkbase stores the labels used to describe concepts. A concept can be associated with multiple labels, for instance labels for different languages. However, it can also be used to add explanatory labels.
- The reference linkbase defines references to external information sources, such as documents, publications, or legal codes.

The many interdependencies can lead to extremely complex taxonomy structures. Fortunately, software is available to check for errors (Daas and Stroom, 2006). Instances can also be checked. Checks of this kind are known as validation. Special XBRL software is available for creating and validating taxonomies and instances (see Section 1.5.3).

The modularity of XBRL means that taxonomies can, and frequently do, import other taxonomies. What is commonly encountered is therefore better described as a Discoverable Taxonomy Set (DTS) rather than a single taxonomy (XBRL, 2006a). As mentioned above, a taxonomy is a combination of an XML schema and associated (one or more) linkbases. Use of the term DTS implies that several linked taxonomies are involved. However, every DTS must have a top-level taxonomy that imports all the others. An instance is created based on the top-level taxonomy, and the instance will accordingly refer to this taxonomy. A DTS facilitates the reuse of data from other taxonomies.

1.5.2 Instance

Every instance must be based on a taxonomy. The taxonomy of an instance determines the concepts that can be reported and their interrelationship and structure. An instance is comprised of a collection of concepts that are defined in the

taxonomy (or DTS), and associated values, which are to be reported on. The values are commonly referred to as facts (XBRL, 2006a). Besides the actual data, the concepts in an instance also have attributes. Attributes specify additional properties of the value of the fact to be reported, such as the context, accuracy and currency, or other unit of a numerical value. All concepts in an XBRL instance must have a context.

A context is composed of data such as the reporting unit and the period over which the instance data reports. It is impossible to interpret instance data correctly in the absence of context.

Numerical values also have information about the accuracy of the value, which is expressed in the 'precision' or 'decimals' attribute. These attributes state the accuracy of the data in terms of the number of significant digits before or after the decimal point (see also: Daas and Stroom, 2006).

If the data in an instance are financial, the instance will contain information on the currency of the fact provided. This is given in the unit. The currency will usually be euros, but it is possible to define custom units.

1.5.3 Software

A limited number of software packages with XBRL functionality are available on the market (XBRL, 2007c). Because XBRL is based on XML, it is often wrongly assumed that standard XML software can be used for XBRL. This is *not* the case, and is a common beginner's error. Only a few manufacturers have developed software specifically for XBRL, which supports the creation and validation of taxonomies and instances (Daas, 2005a). As a means of promotion, software manufacturers often distribute free trial versions of their XBRL packages, allowing users to create and validate XBRL documents for a limited period. This is an excellent way of familiarizing with XBRL. There is additional information about XBRL software on the XBRL.org website (XBRL, 2007c). On the other hand, it is to be expected that, as XBRL becomes more widely used, there will be a corresponding increase in the number of accounting packages with XBRL instance export functionality for financial and other data. The Dutch Taxonomy site provides a list of accounting packages that are in the process of being upgraded with integrated XBRL functionality (NTP, 2007a).

XBRL software can be used to create a taxonomy or instance, but also to verify compliance of these types of files with the XBRL specification. This is also known as validation. XBRL 2.1 is the most recent specification version (XBRL, 2006a). Most XBRL programs support this version. If needed, however, Statistics Netherlands also continues to use and process taxonomies and instances that comply with XBRL 2.0 (XBRL, 2001) as long as this version is still supported by software manufacturers (XBRL, 2007b).

2. Taxonomy

2.1 Short description

The basis of XBRL is the taxonomy. The taxonomy comprises the definitions of the terms to be reported, also known as metadata. Because XBRL is not a reporting standard in a traditional sense (unlike the Generally Accepted Accounting Principles of US-GAAP, or the International Accounting Standards), but simply a technology, multiple taxonomies are created for each region and report type, e.g. based on GAAP or IAS definitions (XBRL, 2007a). Examples include the International Financial Reporting Standards (IFRS), UK-GAAP and the Dutch Taxonomy (NTP, 2007a).

A taxonomy is published on the internet in a form that allows access by XBRL software. The metadata information of an instance (defined in a taxonomy) is thus always available.

2.2 Applicability

Statistics Netherlands uses the following taxonomies:

1) Water boards;
2) Municipalities;
3) Provincial governments;
4) Point of sale system taxonomy (EGS-POS);
5) The Dutch Taxonomy.

In view of its increasing importance, the Dutch Taxonomy is discussed first below, followed briefly by the other taxonomies.

2.3 Detailed description

2.3.1 The Dutch Taxonomy

The Dutch Taxonomy is available for Dutch companies that use XBRL to submit reports to the Dutch government. The taxonomy was created to make it easier and less expensive for companies and governmental organizations to compile and exchange financial data, through a combination of data standardization and ICT (Lammers, 2006; NTP, 2007a). The Dutch Taxonomy primarily addresses the following three financial data types:

1) financial statements for the chambers of commerce;
2) various tax returns (corporate income tax, wage withholding tax, VAT and social security contributions);
3) various data for business statistics.

Alongside Statistics Netherlands, the Dutch Tax office and the Chambers of Commerce are intensively involved in creating the Dutch Taxonomy (NTP, 2007a). The Netherlands Taxonomy Project was set up for this collaboration.

At the time of writing (June 2007), two versions of the Dutch Taxonomy had already been released (NTP, 2007a). The Dutch Taxonomy is actually a collection of multiple DTSes, one for each report. There are four Statistics Netherlands reports: the Structural Business Statistics (SBS), Short-term Statistics (STS), Investment Statistics (Invest.) and Financial Statistics for Small Business (FSSB). For the Short-term Statistics four variant reports are available.

The Dutch Taxonomy comprises several layers (Lammers, 2006), as shown in Figure 1. Starting from the IFRS taxonomy, the Dutch Taxonomy has the following four layers:

1) a common layer with the definitions of all shared concepts and data types;

2) a domain-specific layer with (for Statistics Netherlands) the definitions of all Statistics Netherlands-specific concepts and data types;

3) a formset layer that (for Statistics Netherlands) structures the underlying concepts in common blocks;

4) a report layer, with (for Statistics Netherlands) the taxonomy to be reported comprising the formset files created earlier.

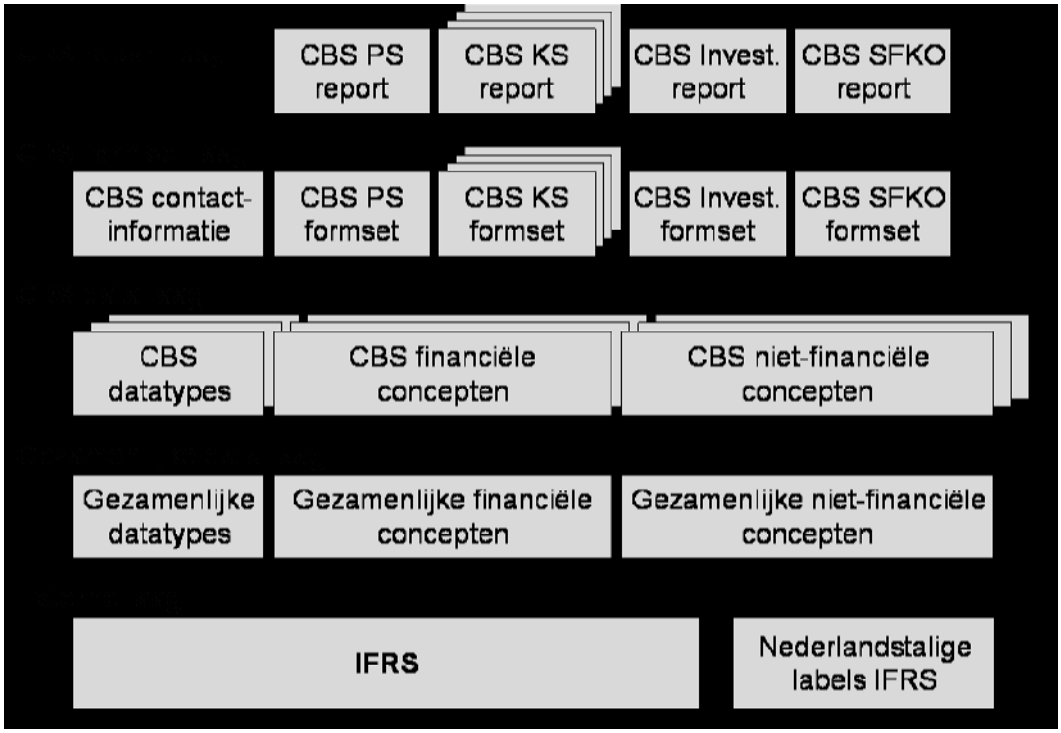
Layers 2, 3 and 4 also occur in the Dutch Tax Office and Chambers of Commerce parts of the Dutch Taxonomy (Roos, 2007). Needless to say, these layers also contain other concepts, data types, formsets and reports.

2.3.1.1 Common part

The Dutch Taxonomy unambiguously records the common metadata of the three parties involved, in what is known as the common data layer, shown in Figure 1. All parties use the common concepts in their reports. The three partners have agreed not only about the definitions of these concepts, but also about the associated XBRL attributes (e.g. data type and period). All common concepts are defined as flow variables (Section 1.5.1). The common part of the Dutch Taxonomy is a first step on the way to a single, government-wide, set of metadata.

The Dutch Taxonomy is based on the international IFRS taxonomy, which is developed and is maintained by the International Accounting Standards Board (IASB). The IFRS accounting standard is used as a basis because it has been mandatory since 2005 for large publicly listed companies in all European Union member states. The IFRS taxonomy is used in the Dutch Taxonomy with the Dutch language label linkbase. This linkbase contains the official Dutch language terms for all IFRS concepts. Metadata that are unique for each partner are stored in that partner's specific domain. Statistics Netherlands has managed the common part of version 1.1 of the Dutch Taxonomy (Roos, 2007).

Figure 1. Overview of the Dutch Taxonomy structure (in Dutch). Each layer consists of one or more taxonomies. Abbreviations are explained in the text.



2.3.1.2 Statistics Netherlands part

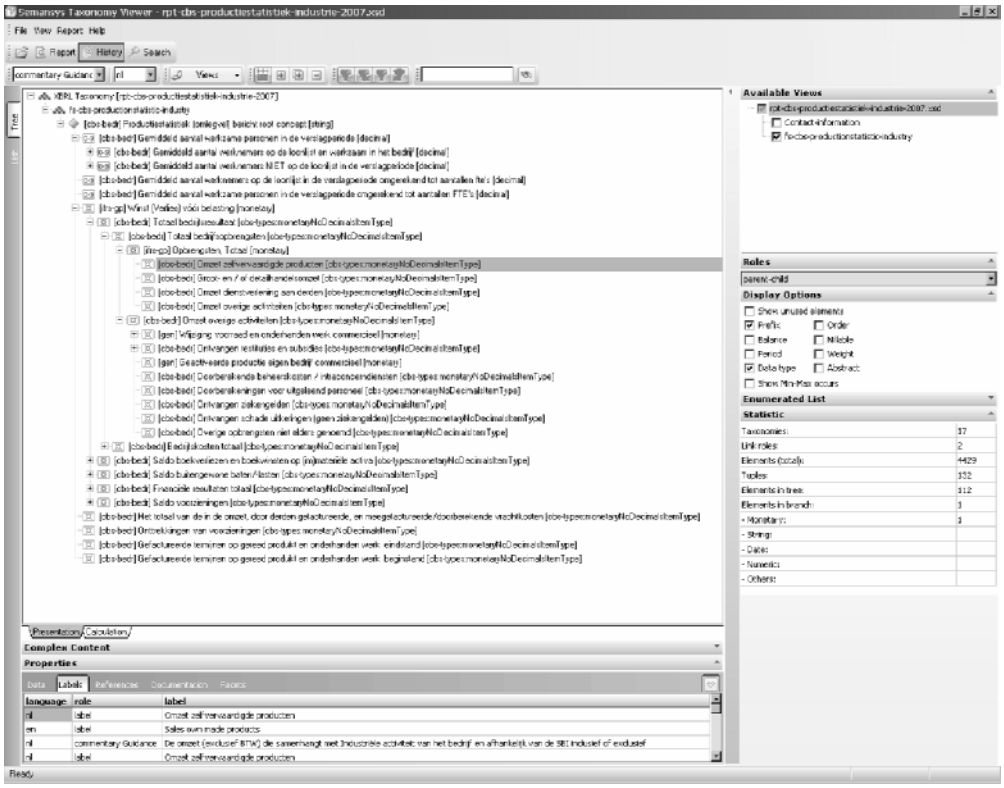
The Statistics Netherlands part of the Dutch Taxonomy consists of multiple layers and four reports (Roos, 2006):

- Structural Business Statistics;
- Short-term Statistics (trade, construction, industry and other);
- Investments;
- Financial Statistics for Small Business (FSSB).

These four reports (DTSes) are based on the questionnaires for each of the statistics concerned. The Statistics Netherlands-specific concepts and data types are part of the Statistics Netherlands data layer of the Dutch Taxonomy (Figure 1). The formset layer defines common data blocks. For the reports mentioned above, there is a formset for each report including the financial data structure plus any explanations and calculations. There is also a formset that defines contact information. The report layer compiles reports by adding the contact information formset to each of the statistical formsets. Statistics Netherlands instances must always be based on the taxonomies in the report layer. An advantage of this kind of layered structure is the efficient reuse of existing concepts and structures and a more easy maintenance. In this way, it is also fairly simple for Statistics Netherlands to add new questionnaires and report variants to the Dutch Taxonomy.

Figure 2 shows the Industry variant of the Structural Business Statistics reportset included in the Dutch Taxonomy version 2.0. In the figure the presentation structure of the concepts to be reported is displayed. The hierarchical structure of the presentation, the data types, and the origin of the various concepts (nl-genbase, ifrs, cbs-bedr) is indicated.

Figure 2. The Industry variant of the Structural Business reportset



2.3.2 Other Statistics Netherlands taxonomies

The best known other taxonomy provided by Statistics Netherlands is the water boards taxonomy (Roos, 2003). Apart from being the first taxonomy to be used in the Netherlands, it was also one of the first to be used by a government body anywhere in the world. The taxonomy was officially adopted in January 2004 and is used by sixteen district water boards for submitting their quarterly financial data to Statistics Netherlands (Roos, 2003). The first version of the taxonomy conformed with XBRL 2.0 specifications, and was later replaced by one compliant with the more recent XBRL 2.1 version. The water board taxonomy gave a substantial boost to the use of XBRL in Statistics Netherlands.

The water board taxonomy consists of two modules (Roos, 2003). The first one contains financial concepts only. There are 651 in total, all of which were defined by Statistics Netherlands. The other module contains the nonfinancial data. It is based on the international Global Common Document taxonomy, which defines a set of 145 concepts for general use, with information about reports, reporting periods and contact information. The water boards need only a limited number of these concepts.

Besides the water boards taxonomy, Statistics Netherlands has also created taxonomies for municipalities and provincial governments. These taxonomies comply with XBRL 2.0 only, and have extremely complex structures. Instances based on these taxonomies involve many contexts to distinguish individual facts, which has hindered creation of a version 2.1 of this taxonomy. The municipalities taxonomy contains two modules; one with financial concepts (101) and one with nonfinancial concepts (2). Approximately 10% of municipalities in the Netherlands deliver quarterly XBRL data to Statistics Netherlands. A similar taxonomy has been created for provincial governments. Three provincial governments deliver quarterly data in XBRL.

The point of sale system taxonomy was created as part of the government EGS-POS programme (In Dutch: Elektronische Gestandaardiseerde Statistiekuitvraag Point of Sale Systemen. Groen, 2004). This programme envisaged the coordinated collection of statistics from the catering and retail trade associations and Statistics Netherlands, in accordance with an XBRL taxonomy. It can be viewed as a forerunner of the Dutch Taxonomy. The supporting functionality was implemented directly into numerous point of sale systems as part of the trial conducted by the participating software houses. Although the trial ended in early 2007, it still produces a regular monthly flow of ten messages. The trial results are discussed in the EGS-POS project final report (GBO.Overheid, 2007).

2.4 Example

2.4.1 Taxonomy creation

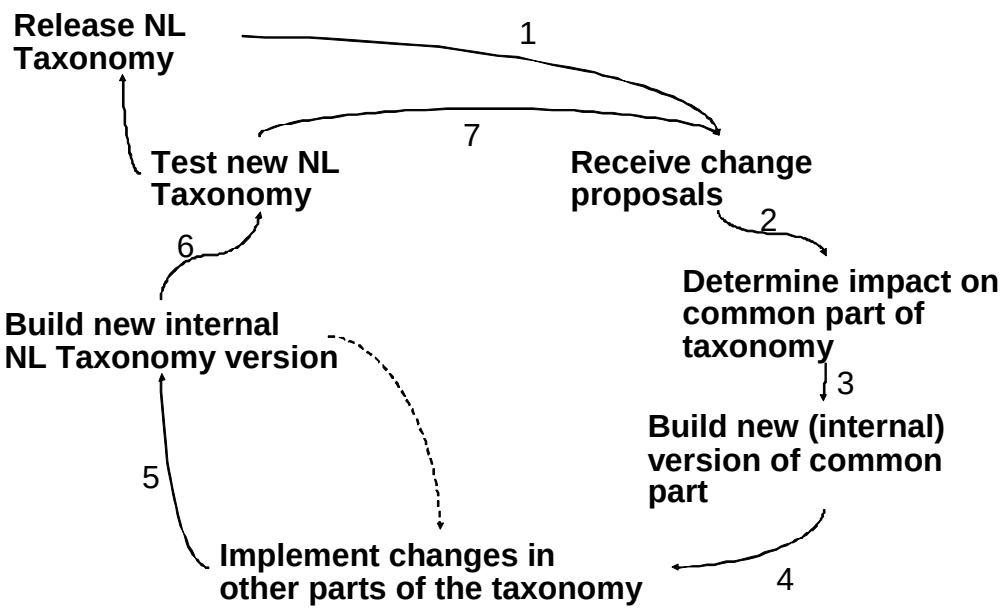
A taxonomy will invariably be built on the basis of existing Statistics Netherlands metadata. This provides the names and data type of the variables. However, this information does not suffice for the creation of an XBRL taxonomy (Daas and Stroom, 2006). Among the additional data needed are the period type and sometimes also the balance type. The requirement to give concepts English names means that these will also have to be added, presumably along with both English and Dutch labels. This is the minimum set of information for creating a simple taxonomy (a schema and a label linkbase). It will be apparent from the above that the metadata of an XBRL taxonomy contains quite some XBRL-specific information. Partly for this reason, Statistics Netherlands does not have a metadata system that contains all this information. This is a subject for further study. The metadata needed are typically supplied by the technical expert concerned in the form of an Excel worksheet, or Access database. After the data have been validated, they can be imported into an XBRL taxonomy creation program, either directly or via a Comma-Separated Value (CSV) format. This process will yield a simple taxonomy. Presentation, calculation and any other relationships laid down in linkbases can be added by using the XBRL taxonomy creation program. Alternatively, the additional data could be imported from an Excel worksheet, or CSV file. After creation, the taxonomy must be validated to confirm that it is free of errors, because it would be of no use otherwise (Daas and Stroom, 2006; Daas and Roos, 2007). The additional requirements for the taxonomies released by Statistics Netherlands are discussed in Sections 2.5 and 2.6.

2.4.2 Taxonomy management

Statistics Netherlands maintains the taxonomies of the water boards, municipalities, and provincial governments. New error-free versions are released on the Statistics Netherlands website. Management and maintenance of the Dutch Taxonomy is far more complex, since multiple organizations are involved in its creation (Roos, 2007). Figure 3 shows the various steps involved in managing the Dutch Taxonomy (Roos, 2007). The Dutch Taxonomy is actually maintained and managed in two separate parts: i) a common part and ii) the partner-specific parts.

The first action is always to identify the consequences of change proposals in the common part of the Dutch Taxonomy (Step 2). A new internal version of the common part is generated after the changes have been implemented, followed by a validation step (Step 3). The version concerned is distributed to all the partners so that each can determine the effect of the consequences of the changes on their part of the Dutch Taxonomy. After investigating, discussing, and resolving any issues, each partner subsequently creates a new version of their part of the Dutch Taxonomy (Step 4). All new parts, including the new common part, are then combined (Step 5). Any problems that emerge at this stage that have consequences for one of the partner-specific areas or the common part are solved in joint consultation. Figure 3 shows this part of the process with a dotted line between the output of Step 5 and the output of Step 4.

Figure 3. The Dutch Taxonomy control and management cycle



The entire new internal version of the Dutch Taxonomy is then systematically tested (Step 6; see also Section 2.5). Provided no problems are found, some necessary final adjustments are needed; such as the inclusion of an unique ‘namespace’ title in the

individual files (NTP, 2007b). Change proposals are raised for any other problems encountered (Step 7). The new version of the Dutch Taxonomy is released on the Netherlands Taxonomy Project website (NTP, 2007a). Statistics Netherlands always creates specimen instances for its own reports in the new Dutch Taxonomy when it is submitted to the NTP for release.

2.5 Characteristics

All Statistics Netherlands taxonomies are validated using at least three XBRL software packages, and *must* be validated flawless by each (Daas and Stroom, 2006; Daas and Roos, 2007). More specifically, no findings may occur on 'error' and as few as possible on 'warning' level. An error and warning free taxonomy complies fully with XBRL specifications. It is our experience that validation with only one software package is no guarantee of full compliance of a taxonomy to the XBRL specifications (Daas and Stroom, 2006).

With the introduction of the Dutch Taxonomy, and following the standard XBRL 2.1 validation, version 2.1 of the Statistics Netherlands part of the Dutch taxonomy will also be validated against the Financial Reporting Taxonomies Architecture (FRTA) specifications (FRTA, 2006). These requirements are of a higher level than the standard XBRL specifications for a taxonomy (Daas and Stroom, 2006). However, the Dutch Taxonomy architecture intentionally deviates from some FRTA requirements (NTP, 2007b).

2.6 Quality indicators

The following three quality indicators and levels are defined for taxonomies.

- XBRL specifications (versions 2.1 and 2.0): 'error' and 'warning' level.

  Minimum requirement for Statistics Netherlands taxonomies: no 'errors' and as few 'warnings' as possible (preferably none).

- FRTA specifications (version 1.0) for XBRL taxonomies (version 2.1): 'must', 'should', and 'may' levels.

  Minimum requirement for Statistics Netherlands taxonomies that are part of the Dutch Taxonomy: as few 'must' and 'should' messages as possible. Exceptions to the above are the intentional deviations from the FRTA requirements, as stated in the architectural specification of the Dutch Taxonomy (NTP, 2007b).

- Architectural requirements of the Dutch Taxonomy: list of intentional deviations from the FRTA standard.

  Statistics Netherlands taxonomies that are part of the Dutch Taxonomy must be fully compliant with the architectural requirements of the Dutch Taxonomy (NTP, 2007b). If they are noncompliant with the FRTA specifications they must comply with the architectural requirements. Software that will also check these requirements is expected to become available in the near future.

3. Instances

3.1 Short description

An XBRL instance is a report containing data. An instance contains the facts to be reported, and is always based on some part of a taxonomy or DTS (XBRL, 2006a). Besides the actual data, an instance also contains additional information such as the reporting period, the currency, and other business or government related items. These data usually resides in the 'context' part of an instance. It is impossible to correctly interpret the data in an instance without context information.

3.2 Applicability

Statistics Netherlands is able to process instances based on the taxonomies given in Section 2.2. On these taxonomies the following seven different instances can be created:

1) Water boards;
2) Municipalities;
3) Provincial governments;
4) Structural Business Statistics (NL Taxonomy);
5) Short-term statistics (NL Taxonomy; 4 variants);
6) Investments (NL Taxonomy);
7) Financial statistics for small business (NL Taxonomy);
8) Point of sale system (EGS-POS.)

3.3 Detailed description

3.3.1 Instance creation

Most financial data that a company reports to the outside world will be available within the organization in some electronic form. Producing an XBRL instance from one of these systems requires (besides the taxonomy) availability of an 'XBRL-enabled' application. This application will be used to link the data in the (multiple) company's accounting system(s) to the concepts defined in the XBRL taxonomy. This is why manufacturers of financial accounting packages are also involved in the Netherlands Taxonomy Project (NTP, 2007).

Apart from using an XBRL-enabled accounting package, it is also possible to create an XBRL instance with specific XBRL software and a taxonomy (Daas, 2005a; Daas and Stroom, 2006). With software of this kind, the data can be keyed in manually, or can be imported from an Excel worksheet or CSV file, for example. An advantage of using dedicated XBRL software is that it can also validate the instance created.

Figure 4 shows an example of part of an instance without using an XBRL tool to retrieve the associated taxonomy information. In the example two concepts are reported: ProfitLossBeforeTax and TotalCompanyResult. Both have a value (284195 and 267000, respectively) and refer to a context ‘c01’ and a unit ‘u01’. The definitions of context ‘c01’ and unit ‘u01’ are also given in the instance. Context c01 concerns ‘Bedrijf A’ (Company A) and the year 2007, and has the status ‘Voorlopig’ (Preliminary). The unit defined refers to the ISO 4217 standard for currency codes and has the value ‘EUR’, for euro’s.

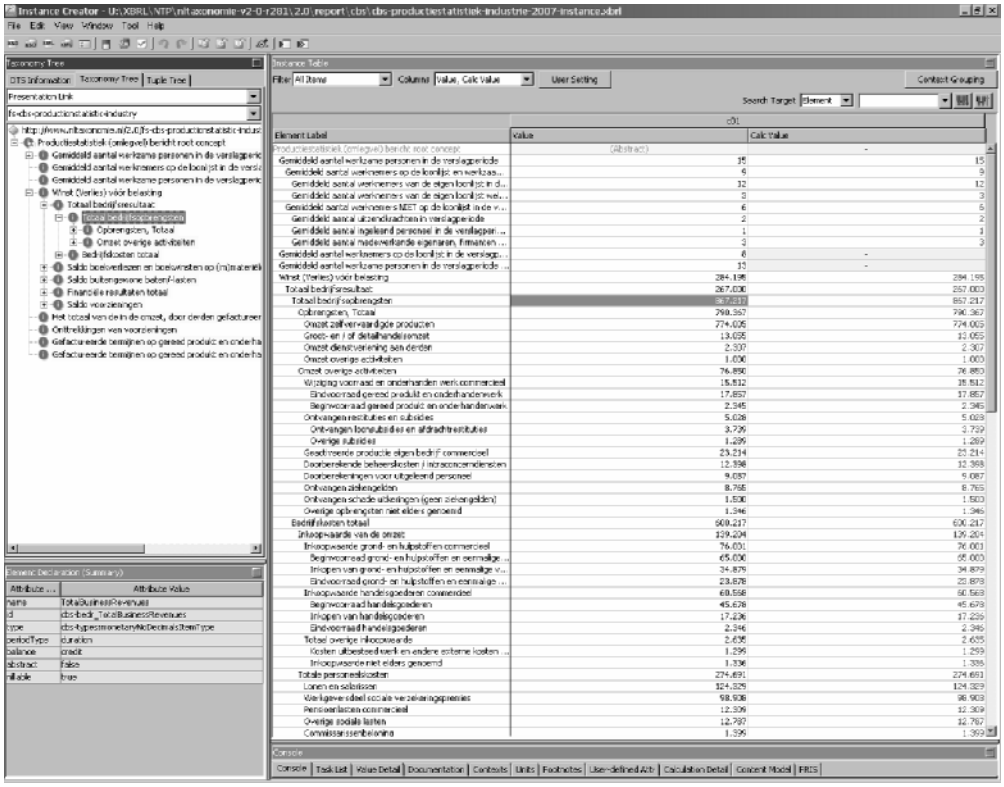
Figure 4. Specimen instance as plain text (without using an XBRL tool)

```
<ifrs-gp:ProfitLossBeforeTax decimals="0" contextRef="c01" unitRef="u01">284195</ifrs-gp:ProfitLossBeforeTax>
<cbs-bedr:TotalCompanyResult decimals="0" contextRef="c01" unitRef="u01">267000</cbs-bedr:TotalCompanyResult>

  <xbrli:context id="c01">
  <xbrli:entity>
    <xbrli:identifier scheme="http://test.cbs.nl">Bedrijf A</xbrli:identifier>
  </xbrli:entity>
  <xbrli:period>
    <xbrli:startDate>2007-01-01</xbrli:startDate>
    <xbrli:endDate>2007-12-31</xbrli:endDate>
  </xbrli:period>
  <xbrli:scenario>
    <cbs-sc:CbsScenario>Voorlopig</cbs-sc:CbsScenario>
  </xbrli:scenario>
</xbrli:context>
<xbrli:unit id="u01">
  <xbrli:measure>iso4217:EUR</xbrli:measure>
</xbrli:unit>
```

Figure 5 shows the same instance, but now as presented with the aid of an XBRL tool (Fujitsu XWand). The taxonomy has been used to add information to the instance document, such as hierarchies and names. Arithmetic checks (calculations) have also been performed on the values in the instance (and were found to be correct).

Figure 5. Specimen instance in an XBRL tool (Fujitsu XWand)



3.3.2 Instance processing

After creation, XBRL instances will be sent to Statistics Netherlands for processing. Submission and processing may follow one of two scenarios (Daas et al., 2007). The first scenario describes the processing of instances based on the taxonomies of the water boards, municipality, and provincial governments. The second one describes the processing of instances based on the Statistics Netherlands part of the Dutch Taxonomy. The following section explains how the two scenarios differ.

3.4 Example

3.4.1 Water board, municipality, and provincial government instances

Instances based on the taxonomies of the water boards, municipalities, or provincial governments are all sent to and processed by Statistics Netherlands in an identical way. This process is shown in Figure 6 (Daas et al., 2007). Routinely, XBRL instances are created, encrypted, and sent as an e-mail attachment to Statistics Netherlands by the reporting parties. First, Statistics Netherlands decrypts the attachment. If the decoded file is identified as an XBRL instance for the water boards, municipalities or provincial governments, an eXtensible Stylesheet Language Transformation (XSLT conversion)[1] is executed. The output of the XSLT instance conversion is such that it can be stored in a Statistics Netherlands database. A separate XSLT is used for each of the instance types (e.g. water boards, municipalities and provincial governments). Instances that cannot be converted in the XLST step generally contain syntax errors. These instances are manually checked and corrected at a later stage. This was a relatively frequent issue in the initial phase. The extensive software support that is currently available to reporting parties has almost eliminated the XLST conversion problem. Disadvantages of using an XSLT are maintenance related and the relative large amount of code needed to convert an XBRL instance.

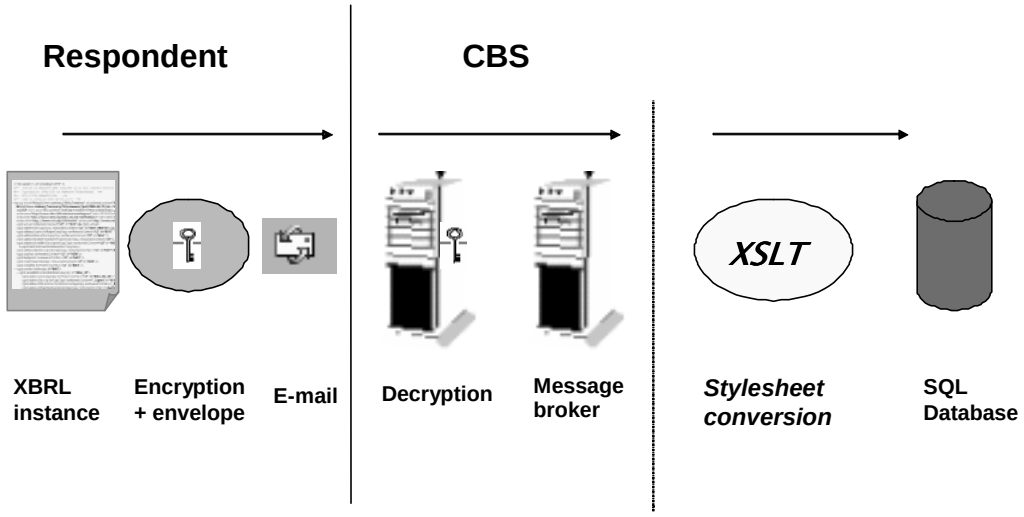
3.4.2 Dutch Taxonomy instances

Instances based on the Dutch Taxonomy are processed in a different way; for these instances a dedicated program is used (Daas and Roels, 2006). The output that the XBRL conversion program produces from the data of the various Dutch Taxonomy instances is in XML, Blaise, CSV, SQL, or Access format. Figure 7 shows the processing of Netherlands Taxonomy instances at Statistics Netherlands (Daas et al., 2007). Instances of this kind are sent via the Governmental GBO web server to Statistics Netherlands. Upon arrival, instances are first inspected to identify the taxonomy/DTS on which the instance is based. Only XBRL instances based on one of the Statistics Netherlands reportsets in the Dutch Taxonomy are used as input to the XBRL conversion program. This program converts the data in the instance in accordance with the settings in the associated configuration file (Daas and Roels,

---

[1] An XSLT is used to convert XML-type files into other XML or text format files (Daas and Roels, 2006).

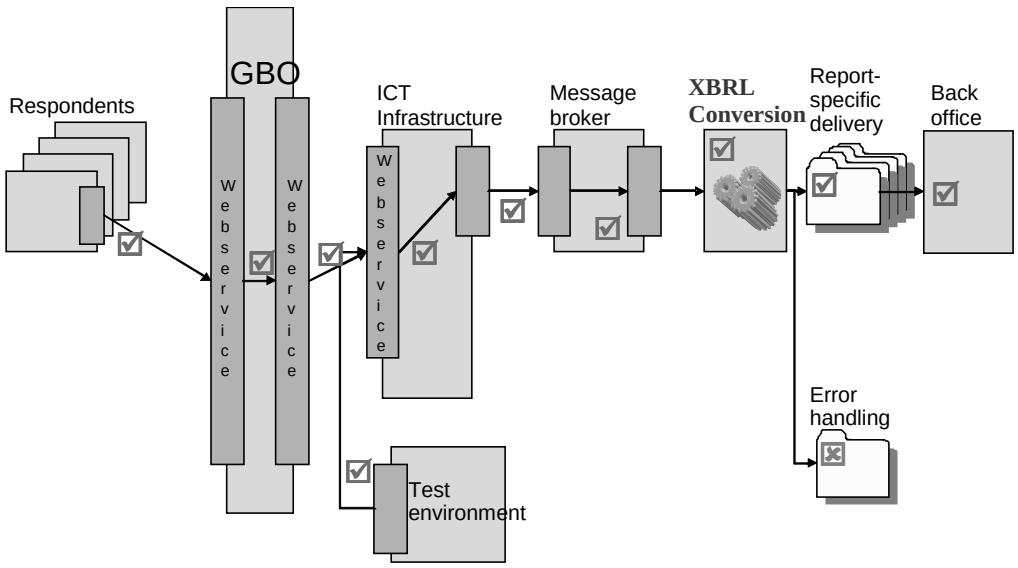
Figure 6. Overview of XBRL processing of the water boards, municipalities, and provincial governments instances at Statistics Netherlands



2006). For each reportset, uniquely identified by the combination of reportset schema name and namespace, settings are defined in the configuration file to exactly specify the actions applied to the data in the instance.

In the settings the output format, the mapping of concept names and variable names, and any additional validations are defined (Daas and Roels, 2006). One of the additional validations included is whether the concept 'CBS userId' is present and has a defined value in an instance. Without this concept there is no way Statistics Netherlands is able to identify the sender of the data. Another validation checks the

Figure 7. Overview of the XBRL instance processing for instances based on the Statistics Netherlands reportsets of the Dutch Taxonomy



presence and value of the two concepts used to report the start and end dates of the reporting period of the instance.

If conversion is successful, the data are stored in a specific folder for each reportset (Figure 7). The back office system of the relevant statistic (SBS, STS, Invest. or FSSB) is able to access the data in the folder to enable further processing. XBRL data based on one of the reportsets in the Statistics Netherlands part of the Dutch Taxonomy are converted at Statistics Netherlands into the most appropriate electronic data format used by the statistic concerned. An advantage of this approach is that the recording of message receipt is processed in accordance with the standard procedure for the statistic concerned.

3.5 Characteristics

An instance sent to Statistics Netherlands should normally be validated by the respondent prior to submission (Daas, 2005b). Upon receipt, however, Statistics Netherlands cannot assume that such a validation has occurred. Experience with processing the water board instances and the quality of the first test instances created by third parties based on the Statistics Netherlands reports in the Dutch Taxonomy have confirmed the need for caution (Roos, 2006). Thorough validation of instances is important and is possible on multiple levels (Daas and Roos, 2007).

3.5.1 XBRL instance validation

Instances can be checked at the technical, syntactic and semantic level. Each of these levels can be performed on XBRL instances (Daas and Stroom, 2006; Daas and Roos, 2007). In essence, the following five clearly different validation levels for XBRL instances can be distinguished:

1) Well-formedness (i.e. an XML syntax check);
2) XML validation relative to the XBRL instance schema definition (version 2.0 or 2.1);
3) XML validation relative to the associated schema (version 2.0 or 2.1);
4) XBRL validation relative to the taxonomy/DTS (version 2.0 or 2.1);
5) Financial Reporting Instance Standards (FRIS) validation relative to the taxonomy/DTS (version 2.1).

The levels are listed above in order of increasing complexity. An XBRL instance that is XML-valid relative to the XBRL instance schema definition (Level 2) will also comply with the requirements of the first level (well-formedness). The use of dedicated XBRL validation software is not necessary for the first three levels of validation. Validation on these levels can be performed i) with free publicly available (i.e. open-source) software, or ii) custom software e.g. using the XML functions of the .NET framework. Dedicated XBRL validation software is required for XBRL taxonomy validation (Level 4) and for FRIS validation (Level 5). In Daas and Roos (2007) additional details of the validation levels are provided. The FRIS

requirements are applicable to XBRL 2.1 instances only, and are of a higher level than the standard XBRL specifications (FRIS, 2004).

3.5.2 Statistics Netherlands instances

Statistics Netherlands routinely adds at least one specimen instance with every taxonomy released. For example, a specimen instance for each reportset and variant of the Statistics Netherlands part of the Dutch Taxonomy is provided on the Netherlands Taxonomy website (NTP, 2007a). All data in the specimen instances are fictional. All the instances must comply with XBRL 2.1 or 2.0 specifications, depending on the XBRL version of the taxonomy. Statistics Netherlands specimen instances for the Dutch Taxonomy must also comply with the FRIS requirements (FRIS, 2004). However, full compliance with the FRIS requirements implies that a taxonomy is FRTA-valid. Since the Dutch Taxonomy intentionally deviates from FRTA (NTP, 2007b), invalid FRTA messages should be ignored in the FRIS validation of the instance.

3.6 Quality indicators

The following quality indicators and levels are defined for instances.

- Well-formedness: yes or no.

  An instance must be well formed or the content will be unreadable by XML/XBRL software. An instance that does not meet this requirement cannot be processed.

- XML-valid relative to the XBRL instance schema definition (version 2.0 or 2.1): yes or no.

  Only the elements and attributes that are defined in the XBRL instance schema are validated. An instance must comply with these specifications, or it cannot be processed.

- XML-valid relative to the associated schema (version 2.0 or 2.1): yes or no.

  An instance must be valid relative to the associated XML schema in its taxonomy/DTS. An instance that meets this requirement has a correct XBRL syntax. It is very likely not possible to process an instance that does not meet this requirement.

- XBRL-valid relative to the taxonomy/DTS (version 2.0 or 2.1): ‘error’ and ‘warning’ levels.

  An instance that produces ‘errors’ on this level is not compliant to XBRL specifications. It is very likely not possible to process such an instance. XBRL instances that only produce ‘warnings’ can very likely be processed. Minimum requirement for Statistics Netherlands specimen instances: no ‘errors’ and as few ‘warnings’ as possible (preferably none).

- FRIS validation relative to the taxonomy/DTS (version 2.1): ‘must’, ‘should’ and ‘may’ levels.

An instance that produces ‘must’ errors is noncompliant with FRIS requirements. It is very likely not possible to process an instance with such errors. Statistics Netherlands specimen instances based on one of the reportsets in the Dutch Taxonomy must produce as few ‘must’ and ‘should’ messages as possible, with the exception of FRTA-related error messages.

4. Conclusion: future developments

XBRL is a relatively new technique that is developing constantly. This chapter lists the latest XBRL developments. All described below refers to XBRL 2.1 specifications (XBRL, 2006a). Three XBRL extensions are described: Dimensions, Formulas and Versioning. Although the use of dimensions is allowed now (XBRL, 2006b), they have yet to be incorporated in Statistics Netherlands taxonomies and the Dutch Taxonomy. The other extensions are -at the time of writing- not yet officially part of the XBRL standard. However, they are expected to be included in it in the near future. This chapter also discusses the consequences of a wider use of XBRL in the statistics process.

4.1 Dimensions

A need has always been felt in the XBRL community to report different aspects, or dimensions, of facts. For example, revenue per product or per country, or costs per department. The dimensional taxonomies specification was drawn up in order to meet this need (XBRL, 2006b). The XBRL 2.1 (and 2.0) specifications have always allowed the use of dimensions (XBRL, 2006a). However, it was impossible to define dimensions explicitly. The dimensional taxonomies specification now enables this option, so that dimensions can be defined and interpreted unambiguously. The dimensional specification also enables defining constraints on the use of dimensions (XBRL, 2006b). In this way, it is possible to permit one part of a business to report data for all regions and to restrict another business to report only on a specific selection of regions.

4.2 Formulas and rules

The type of calculations that can be defined with the calculation linkbase is very limited. Basically only additions and subtractions are possible. With the formula linkbase much more complex calculations can be defined (XBRL, 2006d). The formula linkbase also enables the inclusion of validation rules to be applied to instance data. In addition, the formula linkbase comes with a set of predefined XBRL functions (XBRL, 2006c). These functions can be invoked and executed in a formula or validation rule. A function provides functionality comparable with the complex calculations embedded in standard spreadsheets. The formula linkbase is expected to be added to the XBRL specifications in December 2007.

4.3 Versioning

Reporting rules and legislation change with time. Taxonomies accordingly have to be modified regularly to stay in line with these changes. The XBRL consortium is studying ways of recording taxonomy histories to support this process. Changes will probably be recorded in the form of a linkbase, which is currently referred to as a ‘versioning linkbase’ (XBRL, 2006e). The aim of this linkbase is to store the differences between the ‘old’ and the ‘new’ taxonomies in such a way that the

linkbase file can be used (preferably by a computer program) to deduce exactly which changes have occurred. Another possible function of a versioning linkbase is to support the rendering of various versions of the same instance based on different versions of a similar taxonomy. This function would capture the effects of changes in regulations, and will allow the comparison of financial reports from different years, independent of any changes in the taxonomy.

4.4 Consequences for the statistics process

The developments surrounding XBRL and the Dutch Taxonomy may profoundly change the way in which data on companies are collected, which may affect the statistical process (Roos, 2006). When many companies are in a position to deliver data in XBRL format to the Dutch government, it may happen that these companies will deliver data on larger scale and faster than they do now. Which companies will do so, how often, and which data they will or will not send, is very difficult to predict. What eventually could happen is that a large set of data becomes available via XBRL (perhaps comparable with the size of registers), but that the amount of data will vary -fairly unpredictably- from period to period. If this happens, it will complicate the sampling process, and methods will have to be developed to deal with varying amounts of missing data (see also: Stroks, 2007). Moreover, it will probably also be necessary to cope with differences in the definition of companies on an even larger scale then is the current practice. This necessitates the development of a methodology that includes actual company structures, and a flexible infrastructure is needed to enable the acquirement of supplementary information.

The various possibilities provided by the XBRL 2.1 specification for verifying data quality (e.g. consistency, data types), give Statistics Netherlands and the Dutch government extensive opportunities to ‘outsource’ (part of) the data quality checks to the data supplier or, for example, to a central government system. The opportunities presented by XBRL have yet to be fully utilized. If Dutch companies accept the Dutch Taxonomy as a defining standard and the guidelines for data delivery, XBRL and the Dutch Taxonomy are set to play a major part in the data collection process of Statistics Netherlands.

5. References

Daas, P. (2005a), *XBRL software compared (In Dutch: XBRL programma's vergeleken)*. Internal report, Statistics Netherlands, Heerlen.

Daas, P. (2005b), *Processing of XBRL instances at Statistics Netherlands (In Dutch: Verwerking van XBRL instanties op het CBS)*. Internal report, Statistics Netherlands, Heerlen.

Daas, P. and Roels, J. (2006), *Storage and conversion of XBRL instances (In Dutch: Opslag en conversie van XBRL-instanties)*. Internal report, Statistics Netherlands, Heerlen.

Daas, P. and Roos, M. (2007), *XBRL instance checking and requirements (In Dutch: XBRL instance controle en requirements)*. Draft, Statistics Netherlands, Heerlen.

Daas, P., Roos, M. and Roels, J. (2007), *XBRL instance processing at Statistics Netherlands*. Presentation held at the 15th International XBRL Conference, Developers track session, Munich, Germany (http://conference.xbrl.org/15th/conference/program/by_track/developers#dev06).

Daas, P. and Stroom A. (2006), XBRL 2005: *Taxonomy and instance creation and control*. Statistics Netherlands, Voorburg/Heerlen.

FRIS (2004), *Financial Reporting Instance Standards 1.0*, Public Working Draft. (http://www.xbrl.org/technical/guidance/FRIS-PWD-2004-11-14.htm).

FRTA (2006), *Financial Reporting Taxonomies Architecture 1.0*, Recommendations with corrected errata (http://www.xbrl.org/technical/guidance/FRTA-RECOMMENDATION-2005-04-25+corrected-errata-2006-03-20.rtf).

GBO.Overheid (2007), *EGS-POS operational trial results (In Dutch: Resultaten EGS-POS bedrijfproef)*. Internal report, Statistics Netherlands, Voorburg/Heerlen.

Groen, G. (2004), *EGS-POS project plan (In Dutch: Projectplan EGS-POS)*. Internal report, Statistics Netherlands, Voorburg.

Hoffman, C. (2006), *Financial Reporting Using XBRL: IFRS and US GAAP Edition*, UBMatrix Inc. Lulu Enterprises, Sevilla, Spain.

Lammers, J. (2006), *The Dutch Taxonomy Project and XBRL*. Working Session on Electronic Raw Data Reporting, 6-8 November. United Nations Statistical Commission and Economic Commission for Europe, Conference on European Statisticians, Geneva, Switzerland.

NTP (2007a), Netherlands Taxonomy Project website (http://www.xbrl-ntp.nl/).

NTP (2007b), *Dutch Taxonomy architecture (In Dutch: Architectuur Nederlandse taxonomie)*, Working draft, version 0.6. Netherlands Taxonomy Project.

Roos, M. (2003), *Documentation of XBRL taxonomy, data collection for the water board EMU reports (In Dutch: Documentatie XBRL-taxonomie gegevensaanlevering waterschappen EMU-rapportage)*. Internal report, Statistics Netherlands, Voorburg/Heerlen.

Roos, M. (2006), *Dutch Taxonomy Project summary document. Activities, obstacles and delivery scenarios from a Statistics Netherlands perspective (In Dutch: Overzichtsnotitie Nederlands Taxonomie Project. Activiteiten, knelpunten en aanleverscenario's vanuit een CBS perspectief)*. Internal report, Statistics Netherlands, Heerlen.

Roos, M. (2007), *Managing Taxonomies in a Multi-Party environment, Experiences of Statistics Netherlands in managing the shared part of the Dutch Taxonomy*. Presentation held at the 15th International XBRL Conference, Developers track, Munich, Germany (http://conference.xbrl.org/15th/conference/program/by_track/developers#dev06).

Stroeks, L. (2007), *Data Collection with XBRL for the SBS (In Dutch: XBRL waarneming in de PS)*. Draft, Statistics Netherlands, Heerlen.

XBRL (2001), *Extensible Business Reporting Language (XBRL) 2.0*, Specification, 24 December (http://www.xbrl.org/tr/2001/xbrl-2001-12-14.doc).

XBRL (2006a), *Extensible Business Reporting Language (XBRL) 2.1, Recommendations and corrected errata*, 18 December (http://xbrl.org/Specification/XBRL-RECOMMENDATION-2003-12-31+Corrected-Errata-2006-12-18.htm).

XBRL (2006b), *XBRL Dimension 1.0*, Recommendation, 18 September (http://www.xbrl.org/Specification/XDT-REC-2006-09-18.htm).

XBRL (2006c), *XBRL Functions 1.0*, Public Working Draft, 7 December (http://www.xbrl.org/Specification/XF-PWD-2006-12-07.htm).

XBRL (2006d), *XBRL Formula Specification 1.0*, Public Working Draft, 18 December (http://www.xbrl.org/Specification/Formula-PWD-2006-12-18.rtf).

XBRL (2006e), *Business Requirements for Versioning*, Public Working Draft, 21 February (http://www.xbrl.org/technical/requirements/TVER-REQ-PWD-2006-02-21.htm).

XBRL (2007a), Website XBRL international (http://www.xbrl.org/).

XBRL (2007b), Website XBRL international, XBRL's History (http://www.xbrl.org/ history.aspx).

XBRL (2007c), Website XBRL international, XBRL software tools (http://xbrl.org/tools/).

Version history

Version	Date	Description	Authors	Reviewers
Dutch version: Benaderingsstrategieën / XBRL				
1.0	18-12-2007	First Dutch version	Piet Daas Marko Roos	Vincent de Heij Michel Lindelauf Jean Ritzen
1.1	23-01-2008	Minor modifications to layout	Piet Daas Marko Roos	
English version: Data collection strategy / XBRL				
1.1E	17-02-2011	First English version	Piet Daas Marko Roos	