Nordic Smart Government 2.0
TALTIO Proof-of-Concept
Documentation

Table of Contents

1   Introduction

TALTIO Proof-of-Concept (POC) is the "Test of data warehouse", which was realized in Nordic Smart Government 2.0[1]. The project ownership is within the Finnish Association of Accounting Firms and the Ministry of Transport and Communication (Finland).

"The world's most valuable resource is no longer oil, but data"; the phrase has become known to both the private and the public sector parties and the demand for more quality data is ever growing. The usage of information technology can be perceived as the means to result but there still remains big leaps for our societies to take in order to reach the full potential of big data.

System interoperability is a challenge that affects most businesses and their stakeholders by applications not being able to automatically communicate with each other. Many companies have identified the lunacy in re-keying information from one system to another and built integrations between their systems. As there currently exists no standard interface definitions, these costly custom interfaces is the only way to achieve some level of interoperability between different systems and when we're dealing with point-to-point integrations, the number of interfaces grows exponentially in relation to the amount of systems in use.

Besides being very costly, the custom interfaces often also mean loss of data and risk of lowering the data quality as manual work is always involved in the creation of such customized interfaces. As the value of quality data increases it is these factors that become even more important for obtaining a standard for transactional business data.

---

[1] http://www.nordicinnovation.org/news/innovating-interaction-between-nordic-companies-and-governments/
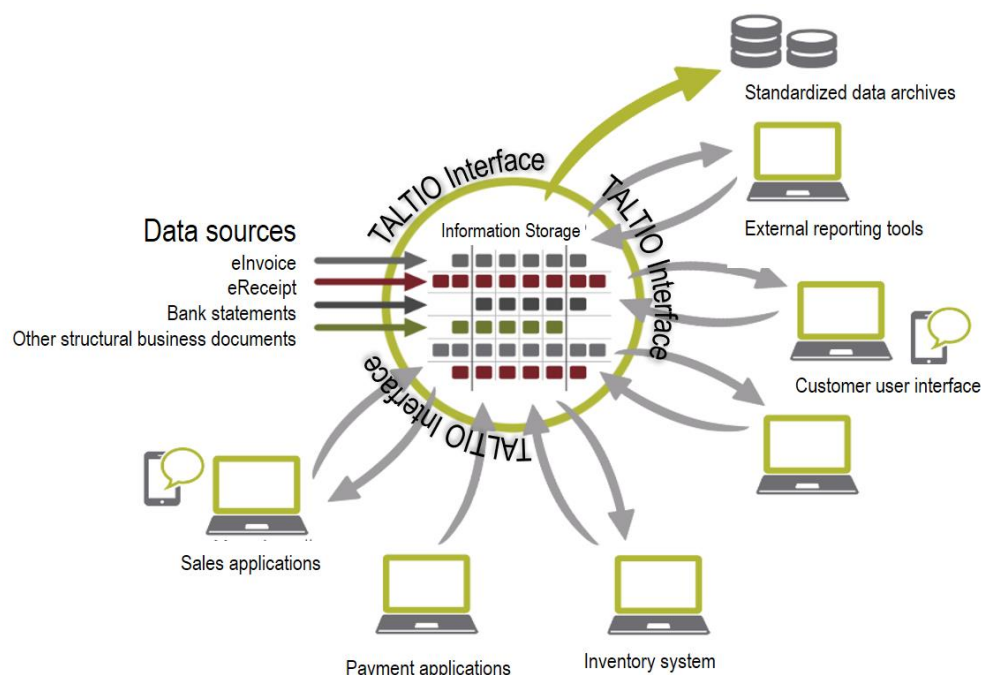
Figure 1.1. TALTIO as a standard interface enabling interoperability between systems

Figure 1.1. above illustrates the vision of future ERP systems fully interoperable enabled by a common interface and orchestration. In a digital interoperable infrastructure the data flows automatically with-in the information systems and also in to regulatory reports in real time which enables the maximum potential of the utilization of the data. As more and more processes can be digitized, unproductive manual tasks can also be eliminated.

The standardized interface is also the key to acquire more details on the business transactions such as product information and other measurable details. More quality data enables new innovations for data analytics for both private and public sector needs. More over the interoperability will lead to reduced transaction costs, increased efficiency, more transparent procedures and better communication.

The XBRL Global Ledger Taxonomy Framework (XBRL GL) is a holistic and generic XML and XBRL-based representation of the detailed data that can be found in accounting and operational systems, and is meant to be the bridge from transactional standards to reporting standards, integrating the Business Reporting Supply Chain.

XBRL GL can be used by computer programs for information interchange of accounting General ledger balances (summarized information) as well as complete accounting ledgers (payables, receivables, inventory, payroll, order entry, purchasing, banking) supporting object oriented accounting, quantity accounting and transparency support.

In this proof-of-concept the XBRL GL is leveraged with minor local extensions including Finnish labels (later referred as TALTIO) and it is used as a common representation of all kind of business documents (invoices, cash receipts, bank account statements etc.). Having all accounting source data in a structural, standardized format with complete business data enables the automation of account posting and regulatory reporting to name a few.

2   Scope

This is a documentation of the implementation process and demo of the TALTIO POC service. The aim is to provide means to realize this POC in another environment and to provide insight for future development in order to achieve automated accounting and external reporting based on structural business data. In the documentation we describe the TALTIO data warehouse covering the essential implementations, enabling factors, challenges and found solutions.

The scope of the POC is to prove that the automation of accounting related processes such as entry posting and external reporting are possible and that more quality and real-life data could be captured from the business processes with less manual work when we rely on standards and structural data.

The TALTIO interface is widely used in the POC as it enables the automation of processes selected here, however, this documentation does not take a stance on the orchestration, connectivity, authorization or data security related issues. These matters are perceived to be determined on the next interoperability tire.

The chapter three provides a recommended reading list that might give broader technical insight to the reader especially if one desires to make use of the technical implementations produced in the POC. The nature of this documentation is not technical and therefore prior knowledge about the below-mentioned technologies is not assumed from the reader.

In chapter four a short glossary is provided to clarify usage of some terms that might otherwise seem confusing due to their technical nature. Chapter five summarizes the data tools used in the POC whilst chapters six and seven document the actual implementations. In chapter eight a walkthrough of the POC demo is provided with references to the demo materials. Finally in chapter nine a summary and key findings is presented together with chapter ten's conclusions and suggestions for further development.

3   Recommended reading

3.1   W3C

XLink: XML Linking Language (XLink) Version 1.1[2]

XML: Extensible Markup Language (XML) 1.0[3]

XSLT: Extensible Stylesheet Language Transformations 2.0[4]

3.2   XBRL

XBRL 2.1: XBRL Specification[5]

XBRL Dimensions 1.0: XBRL Dimensions[6]

Registries. Units Registry 1.0: Structure, Data Type Registry 1.0: Structure, Data Type Registry 1.1[7]

XBRL Global Ledger[8]

3.3   ISO 20022

ISO 20022: Universal financial industry message scheme[9]

3.4   CEN: eInvoicing Semantic Data Model

EN 16931-1:2017 Electronic invoicing - Part 1: Semantic data model of the core elements of an electronic invoice[10]

4   Glossary

XBRL Taxonomy
A set of XML Schema and linkbase definitions that combine into a complete set containing one or more entry point schemas that reflect the reports that are

---

[2] https://www.w3.org/TR/xlink11/
[3] https://www.w3.org/XML/
[4] https://www.w3.org/TR/xslt20/
[5] https://specifications.xbrl.org/spec-group-index-group-base-spec.html
[6] https://specifications.xbrl.org/spec-group-index-group-dimensions.html
[7] https://specifications.xbrl.org/spec-group-index-registries.html
[8] https://specifications.xbrl.org/spec-group-index-xbrl-gl.html
[9] https://www.iso20022.org/about_iso20022.page
[10] https://standards.cen.eu/dyn/www/f?p=204:110:0::::FSP_PROJECT:60602&cs=1B61B766636F9FB34B7DBD72CE9026C72

included in that taxonomy. A taxonomy combines from one or multiple modules that can be perceived as logical sub-parts of the taxonomy.

Taxonomy user
Taxonomy user refers to a person or a party using a taxonomy for;
1. data representation/reporting purposes,
2. building extensions or
3. applying the taxonomy in other use cases such as software development.

XBRL Instance document
An XML document that is constructed according to some entry point schema in the taxonomy.

Inline XBRL (iXBRL) Instance document
An (X)HTML document that is constructed according to some entry point schema in the taxonomy.

Preparer
A person or an organization preparing the reports/instance documents according to the taxonomy.

5   Data tools

5.1   SBR code set

The SBR code set can be perceived as a referential standardized chart of accounts (Attachment 1). The XBRL and Tax report mappings attached to the SBR code set (also available as machine readable format))are the enablers of the report automation.

5.2   Taxonomies

There are two main XBRL taxonomies that were used in the TALTIO POC solutions:

- TALTIO: the Finnish extension of the XBRL GL taxonomy (Attachment 2: taxonomies/XBRL-GL-PWD-2016-12-01-fi-2017-12-01) and

- The Finnish SBR-taxonomy for financial statement, tax and statistics reporting (Attachment 2: taxonomies/fi-sbr-01-01-2016_pub).

5.3   Mappings

- Mappings between TALTIO and other business documents (Attachment 3: Attachment3_TALTIO_definition_and_mappings.xslx) containing:

  ➢ TALTIO - Finvoice 2.01

  ➢ TALTIO - eReceipt

  ➢ TALTIO - Teapps 2.7.2

> ➢ TALTIO - UBL 2.1
>
> ➢ TALTIO - Bank To Customer Statement V02 camt.053.001.02

- Mapping between SBR code set and the Finnish SBR taxonomy (provided in TALTIO (XBRL GL) format using the Summary Reporting Contextual Data (SRCD)-module (Attachment 4: Attachment4_SBR_codes_XBRL_mapping.xml)

5.4   XSLT stylesheets

XSLT stylesheets provide means to transform structural data from one format into another by specifying a set of transformation rules.

In the TALTIO project a set of XSLT stylesheets was produced based on the mappings between TALTIO and other core business document formats; eInvoices, eReceipt and bank account statement (Attachment 2: Source_data_and_XSLT\XSLT_stylesheets).

5.5   Arelle

Arelle is a license free and open source platform for XBRL. The tool can be used for viewing, creating and validating XBRL data in a human readable way.[11]

6   TALTIO Data Warehouse

This chapters provides documentation on the technical solutions, content and access information.

6.1   TALTIO POC implementation

The TALTIO POC is based on TALTIO Web Resource implemented in Tieto Business Information Exchange (BIX) cloud service provisioning infrastructure, a demo web application and a commercial XBRL compliant package XKUBED. The TALTIO Web Resource contains financial information management supporting semantics defined in the TALTIO specification, harmonization and uploading of new information from the input data sources and the TALTIO Web Resource exposure using REST[12] web services with XBRL GL and JSON[13] data formats.

The XBRL GL based TALTIO interface described in chapter 6.4 is ideal for reporting purposes and for exchanging financial information between services and systems. The JSON data format option enables agile development of web and mobile applications, which typically need only a subset of TALTIO data elements.

---

[11] http://arelle.org/
[12] https://en.wikipedia.org/wiki/Representational_state_transfer
[13] https://en.wikipedia.org/wiki/JSON

6.2   Data management

Data management in the TALTIO Web Resource is based on MongoDB[14] (*from hu**mongo***us*), which is a free and open-source cross-platform document-oriented database program. MongoDB is available at no cost under the GNU Affero General Public License, version 3 and the language drivers are available under an Apache License. In addition to those MongoDB Inc. offers proprietary licenses also.

MongoDB is a distributed database at its core, so high availability, horizontal scaling, load balancing and geographic distribution are built in and easy to use. Its design philosophy is focused on combining the critical capabilities of relational databases with the innovations of NoSQL technologies. MongoDB is classified as a NoSQL database and its structure is based on JSON-like documents with schemas. It provides powerful ways to access and analyze data with ad hoc queries, indexing and real time aggregation.

6.3   Input data sources

In order to facilitate the automation of the accounting process, it is expected that the majority of the accounting source data may be found in structural electronic format that is supported by the transformation tools provided. However it is recognized that not all source data is available in structural format and for those entries manual input will be necessary.

For the following data formats there exists a mapping to TALTIO format and readymade XSLT transformation stylesheets (provided in Attachment 2) that are able to convert data from the source data formats into TALTIO format.

E-invoice:

- Finvoice 2.01[15]
- TeappsXML 2.7.2[16]
- UBL 2.1[17]

E-receipt:

- Finvoice 2.01[18]

Account statement:

- ISO20022 - Bank To Customer Statement V02 camt.053.001.02[19]

6.4   TALTIO interface

The TALTIO interface is a data format specification described in the XBRL Global Ledger (XBRL GL)[20] taxonomy for its core parts. XBRL GL was developed by XBRL

---

[14] https://www.mongodb.com/
[15] http://www.finanssiala.fi/finvoice/Sivut/default.aspx
[16] https://www.tieto.com/services/business-process-services/business-information-exchange/tieto-e-invoicing/teappsxml-descriptions-tieto-invoice-messages
[17] https://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ubl
[18] http://www.finanssiala.fi/finvoice/Sivut/default.aspx
[19] https://www.iso20022.org/standardsrepository/public/wqt/Description/mx/camt.053.001.02.
[20] https://www.xbrl.org/the-standard/what/global-ledger/

International which is a global standards development organisation that exists to improve the accountability and transparency of business performance globally, by providing the open data exchange standards for business reporting.

TALTIO is a Finnish national extension to the base XBRL GL taxonomy adding a handful elements that were perceived as necessary in order to capture certain level of data. The TALTIO taxonomy was originally created as a joint effort of the Finnish Association of Accounting Firms and XBRL Finland consortium[21]. In the TALTIO project the taxonomy was developed further alongside with TALTIO mappings and XSLT[22]-transformers.

The entire TALTIO taxonomy contains around 400 element definitions which are not all used in the POC. In annex A there is a listing of the element descriptions that are used in the POC implementation. The taxonomy definitions may be found in the Attachment 2.

7   User Interface Functionalities

7.1   Data upload

Data that is obtainable in one of the supported data formats may be converted into TALTIO format automatically by using the provided XSLT transformers. These source data may include information regarding the following business transactions:
- Bank statements:
    - Payments to sales invoices/receivables
    - Payments to purchase invoices/payables
    - Bank charges
        - loan repayment/payment
        - transaction charges/fees
    - Card payments charges
- Purchase orders / invoices
- Sales Orders / invoices
- eReceipt of purchases or costs

7.2   Controls and automation to the input data (not implemented in POC)

The input file has to fulfil the minimum requirements as to the content and quality defined for each type of business document declared as input data source in their corresponding documentations.

Before the transfer in to the database, some additional information may be added to the document. This might be for instance bookkeeping account numbers and also

---

[21] http://www.tieke.fi/display/XBRL/XBRL+Suomi
[22] https://www.w3.org/TR/xslt

SBR code references. The data may be added by some automated process or by manual input.

In real use there should be a front system that would correspond to invoice approval system ensuring that the data is qualified to be entered as accounting entries in the data base. In the TALTIO environment these kind of data quality assurance procedures may also be harmonized or reused as there is a number of validation tools applicable for XML/XBRL based data such as Schematron[23] and XBRL Formula[24].

7.3   Manual data entry

In the future when all business documents are structural and in a standardized format there shall be very minimal need for manual data entries but in the case of the POC such cases are identified as:

- transaction summary from subsystem (payroll, inventory);
- inventory entries goods received (purchases), when used also for inventory purpose;
- inventory entries goods delivered (sales), when used also for inventory purpose;
- adding an internal accounting (bookkeeping) voucher information (depreciations or adjustments);
- invoices or receipts on paper.

Therefore a manual data entry possibility is also produced as part of the TALTIO POC.

7.4   Data export and reporting based on TALTIO interface

The TALTIO POC database shall export all entries in the TALTIO format for the chosen financial period. All accounting entries must refer to a valid standard business reporting code and contain all other necessary information required in the external business reporting process. In fact the exported TALTIO data contains much more information that the user interface is able to show due to the normalization of one-to-many data structures in the view plane. The actual data is in its original TALTIO form all along in the data base.

External reporting tool XKUBED was selected to be used in the POC. Other XBRL GL compliant tools may also be used. XKUBED is able to aggregate XBRL GL data and form other taxonomy's defined XBRL reports based on a special mapping file that must be provided for the tool.

---

[23] http://schematron.com/
[24] https://www.xbrl.org/wgn/xbrl-formula-overview/pwd-2011-12-21/xbrl-formula-overview-wgn-pwd-2011-12-21.html

The mapping (see chapter 5.3) defines which accounts in relation to other data (such as subaccount) should be aggregated into which element in the target XBRL report (e.g. financial statement or tax return). The target XBRL report shall be the "extended financial report" for limited companies defined in the Finnish SBR taxonomy.[25]

7.5   General journal and general ledger

The general journal and general ledger are views on the financial period's accounting entries. The general journal is presenting the entries in the timely order they have occurred whereas the general ledger presents the entries grouped by their bookkeeping account. The general ledger can be perceived as bases for the creation of the trial balance as well.

From a technical perspective both general journal and general ledger are a combination of sorting and grouping functions applied to the set of entries.

7.6   Invoice reconciliation

The process of invoice reconciliation matches the information on the bank account statement to the invoices the company has issued or received. The output of this process is a listing of open receivables and open payables.

Listing of open receivables

Search for all entries from the sales journal with row type 'header' and match the *Original document identifier code* with entries from the giro journal (bank statement) or from standard entries where row type is 'row'. For matches found, sum the amounts of the entries. In case of sum other than zero or no matching giro entries found, list as an open receivable.

Listing of open payables

Search for all entries from the purchases journal with row type 'header' and match the *Original document identifier code* with entries from the giro journal (bank statement) or from standard entries where row type is 'row'. For matches found, sum the amounts of the entries. In case of sum other than zero or no matching giro entries found, list as an open payable.

The invoice reconciliation process is simple to define but challenging to reliably implement because of incomplete source data. In particular misuse of the payment reference (*Original document identifier code* in TALTIO) in invoices is causing issues as sometimes invoicing parties may store some other data in this dedicated field and there are no validation procedures in place for this data.

Also for cash receipts the reconciliation is challenging as no payment reference is given. In the future as the eReceipts gain more familiarity this might change as the

---

[25] The Finnish SBR taxonomy: https://www.tieke.fi/display/XBRL/SBR+Taksonomia

bank statements are going to include the eReceipt's archiving reference and the payment may be matched by this identifier.

In order to avoid manual intervention, additional rules must be defined. For the POC the purchase invoice reconciliation process is additionally:

- Searching entries from purchase journal where row type is 'header' and from giro journal where row type is 'row' where the announced *Original document identifier code* (payment reference) matches to the *Document number* (invoice number) in the purchase journal entry. For matches found, sum the amounts of the entries. In case of sum other than zero or no matching giro entries found, list as an open payable.

- Searching entries from purchase journal where row type is 'header' and from giro journal where row type is 'row' where the entries' *amount* and *Identifier description* (trade partner name) match. For matches found, sum the amounts of the entries. In case of sum other than zero or no matching giro entries found, list as an open payable.

Reconciliation of eReceipts

Traditionally in the current accounting systems cash receipts are attached to the corresponding payment information in the bank statement. This is generally done by matching the seller details, paid amount and purchase date between these documents.

In the future once eReceipts are available, the reconciliation process is also possible to automate. As the merchandise is already paid for, no payment reference is available but an archiving reference can be stored to the receipt. The same archiving reference may be passed on with the payment and thus be outputted also onto the bank statement which will then be the basis for the reconciliation.

7.7 Automated account posting

In the ideal case the accounting information is already attached to the purchase order process but in reality this data is not present in many invoices. Automated account posting can be realized through AI, machine learning or setting up a base of assumption rules. The AI and machine learning approach requires that the application is able to understand the meaning of the chart of accounts and other classifications used in cost pools and the likes.

By standardizing the (referential) chart of accounts and other classifications used in the accounting and reporting processes within the company, it is possible to produce base definitions that can be made use of in all systems for all parties involved; trading partners, regulators, analysts and so on.

The underlying task of account posting is to take a stand on different kind of attributes regarding the business transaction at hand. Many of these attributes

derive from regulatory reporting requirements and ultimately from the accounting and tax laws. These reporting requirements are also broken down as the Standard business reporting (SBR) code set which can be perceived also as a referential chart of accounts. For further reading the SBR code set v2015 with Swedish labels and mappings to tax reports and SBR taxonomy can be found in Attachment 1 (excel file).

Although the current representation of the SBR code set is in a form of a list, it may also be perceived as a dimensional data model (see example below).
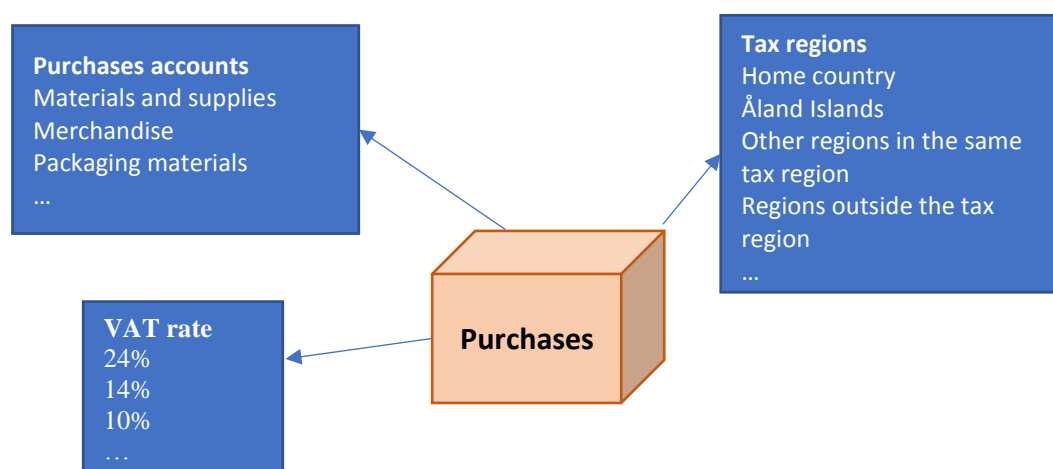
Figure 7.7.1. Multi-dimensional financial data in accounting

In most cases the deduction of account posting in relation to the defined attributes can be automated based only on the information on the invoice. As an example the attributes given in figure 7.7.1:

Purchases account category may be derived by retrieving product info by product code (e.g. EAN code) in the invoice. The tax region may be deducted by seller's country code or VAT number.

The balance sheet counter entry account posting rules are summarized in table x in relation to the SBR code set. The only built in logic in the TALTIO POC database is that there are three different categories for entries related to one document:

- Header: entries are actual total amounts of cash received or paid (invoiced);

- Header VAT: entries are VAT totals for the document;

- Row: entries are net amounts of purchases/sales.

The calculation for VAT reports may be deducted by the following information:

- domestic VAT: for each Header VAT type entry summing per each tax rate and tax amount, where tax code is 'S' (standard tax handling)

- reversed VAT liability: calculation based on each Row type entry where tax code is other than 'S' and the net amount on that entry

Table 7.7.1. Standard default balance sheet account posting in relation to the Finnish SBR code set.

Document type	Header	Header VAT		
		24 %	**14 %**	**10 %**
Purchase invoice	29060 (Trade creditors, short term)	1762 (VAT receivables, short term)	1762 (VAT receivables, short term)	1762 (VAT receivables, short term)
Sales invoice	1700 (Trade debtors, short term)	292041 (General sales VAT debt, short term)	292042 (First reduced sales VAT debt, short term)	292043 (Second reduced sales VAT debt, short term)
Receipt (purchase)	19191 (Cash at bank, money transfers intermediate accounts reconciliations for own use)	1762 (VAT receivables, short term)	1762 (VAT receivables, short term)	1762 (VAT receivables, short term)
Bank account statement	19100 (Cash at bank A)			

7.8 Correction posting

At times there are entries that require corrections to be made. Most commonly these occur in the bank statement entries as bank charges that are automatically deducted from the account without an official invoice being sent. The correction posting functionality in the TALTIO POC is defined as follows:

1. The original entry line is set to 'cancelled' (Memo line)

2. A copy of the original entry is created but with all amounts (monetary and products) multiplied by (-1). This entry is also set to state 'cancelled' (Memo line)

3. Another copy of original entry is created (amount as in the original) but with possibility to edit other column values in the user interface. This entry remains in state 'active'.

7.9   Calculation of trial balance (not implemented in the POC)

A trial balance is a bookkeeping or accounting report that lists the balances in each general ledger account. In the case of the POC, the Finnish SBR code set (referential chart of accounts) is the chart of accounts used in the trial balance export (see chapter 7.4. about data exporting).

In accounting systems the profit/loss for the period is a computational figure that is generated as a sum of all income and expenditure and then shown as the last row on the income statement "Profit (loss) for the period" and also in the balance sheet as an item of equity (see figure 7.9.1).

Operating margin	**-3,067.5**
Taxes and state subsidies	
Tax income	3,205.6
State subsidies	316.3
	3,521.9
Financing income and expenses	
Interest income	82.8
Other financing income	50.6
Interest expenses	-18.0
Other financing expenses	-0.4
	114.9
Annual contribution margin	**569.4**
Depreciation, amortisation and impairment	
Planned depreciation	-337.2
Amortisation and impairment	0.0
	-337.2
Extraordinary items	
Extraordinary income	239.5
Extraordinary expenses	-1.5
	238.0
Profit for the financial period	**470.2**
Increase (-) or decrease (+) in depreciation difference	-13.9
Increase (-) or decrease (+) in reserves	16.3
Increase (-) or decrease (+) in funds	-7.5
	-5.1
Surplus / deficit (-) for financial period	**465.0**

EQUITY AND LIABILITIES (EUR million)	**31 December 2016**
Equity	
Initial capital	2,972.4
Revaluation reserve	1,806.4
Other internal funds	522.3
Other shareholders' equity	729.0
Surplus / deficit (-) from previous financial periods	4,479.3
Surplus / deficit (-) for the financial period	465.0
	10,974.4

Figure 7.9.1: Profit (loss) shown in the income statement and as an item of equity

In the TALTIO POC such computational functionality to calculate the profit or loss for the period was not feasible to create with in the time frame although it's very simple to define (sum of all entries having greater or equal account id than '3000' where the income statement accounts begin). Instead this figure was created as a manual entry as its presence in the data export for trial balance is crucial (otherwise balance sheet wont balance). The trial balance is then used as a base for the financial statement and tax return to be created from it.
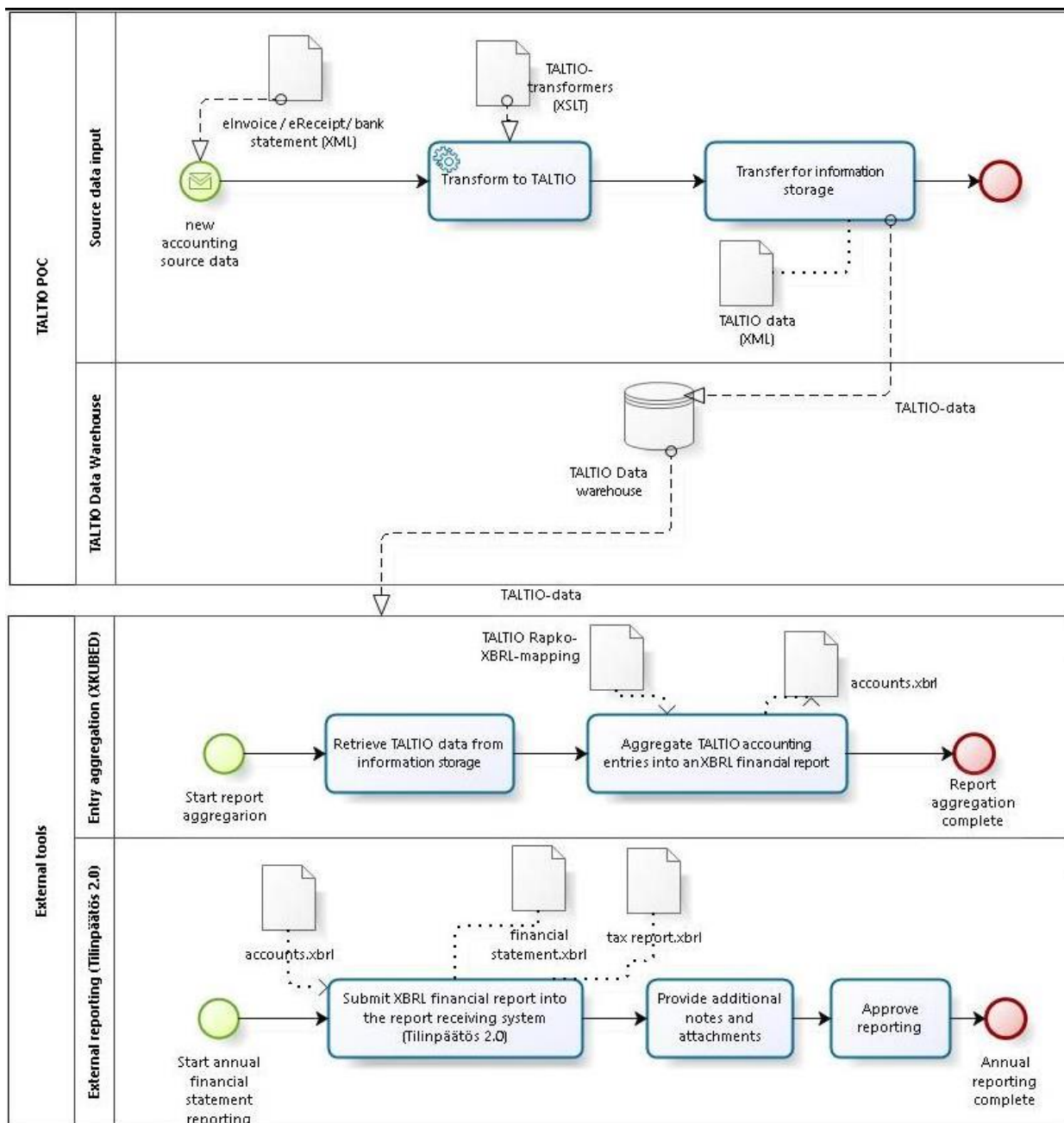
8   Demo

Figure 8.1. TALTIO POC demo

In this chapter a walkthrough of the TALTIO POC demo is provided show casing the implemented functionalities as shown in figure 8.1. In the figure there are two pools; "TALTIO POC" which is presenting actual products of the POC and "External tools" which demonstrates the usage of other tools to realize the regulatory reporting. All the documents referred here are available in the Attachment 2 "Demo_material" (zip file).

The process starts with the event of a new business transaction and its trailing voucher; cash receipt, invoice or bank account statement. The user provides the vouchers in one of the supported data format and then the business documents

may be transformed into TALTIO format with the standard XSLT transformers provided. Quality checks may be performed by using one of the designated validation technologies such as XBRL Formula[26] or Schematron[27] but this functionality was not implemented as a part of the POC. Once in TALTIO format, the data may be uploaded to the to the data warehouse.

In the case of POC all the source data is uploaded manually to the system but in a real production environment different kind of applications such as POS system, banking system or inventory management system would be integrated to the data warehouse via the TALTIO interface and data would flow automatically.

The data warehouse itself doesn't perform any operations to the data. The user interface also implemented in the POC provides different kind of data querying functionalities, manual entry posting and data upload and download functionalities. These are demonstrated in section 7.3.

The complete set of accounting entries for the financial period is given as input for the external reporting process. An external tool[28] shall be used in order to create the financial report in XBRL[29]. In the demo the data is uploaded to the tool manually but system-to-system integrations are also possible in which case this part of the process would be happening behind the scene. Finally the XBRL financial report is submitted to regulatory report receiving system.

8.1   Source data input

By default the entries are automatically created and stored in the TALTIO format that the TALTIO data warehouse is capable to intake. New data can be uploaded to the data warehouse by the "Add to database" functionality which opens a simple file uploading window.
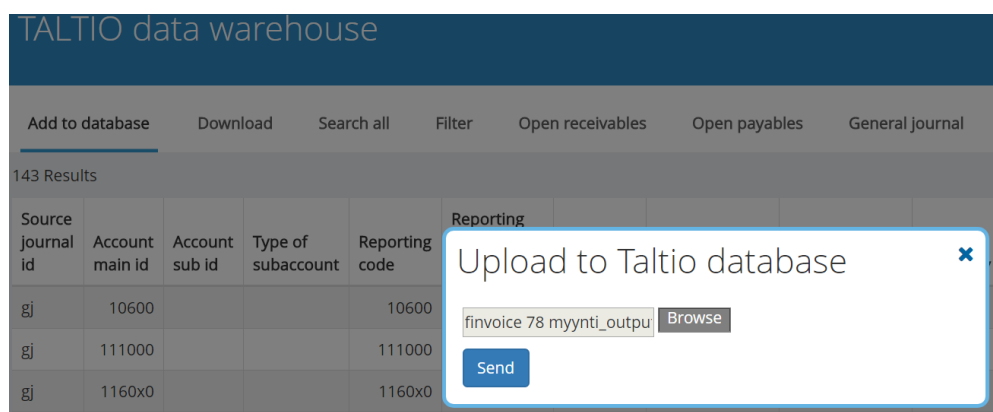

Figure 8.1.1. Uploading new data to the data warehouse

---

[26] http://www.xbrl.org/specification/formula/rec-2009-06-22/formula-rec-2009-06-22.html
[27] http://schematron.com/
[28] http://xkubed.com/
[29] The external XBRL reporting taxonomy chosen is the Finnish SBR taxonomy

8.1.1   Automated entries creation

In the attachment 2 "Demo materials" (zip file) there is a subfolder "Source_data_and_XSLT" contains a set of input data (business documents) and readymade XSLT transformation style sheets that may be used to create TALTIO data. The folder structure contents is as follows:

- Source data (original business documents):
    - Bank account statements (each XML file contains account entries for one day)
    - Purchase invoices
    - Sales invoices
    - Cash receipts
    - Opening balances (ending balances of the previous fiscal period)
- XSLT transformation stylesheets: For each corresponding business document type (apart from opening balances) there is a dedicated XSLT stylesheet
    - Camt_053-TALTIO: For bank account statements complying with ISO20022 Camt 053 BankToCustomerStatement[30]
    - Finvoice-TALTIO: For invoices complying with Finvoice 2.01[31] specification
    - Teapps-TALTIO: For invoices complying with Teapps 2.7.2 specification[32]
    - UBL-TALTIO: For invoices complying with UBL 2.1 Invoice specification[33]
    - eReceipt-TALTIO: For eReceipts complying with Finvoice[34]
- Converted TALTIO data (the source data converted into TALTIO data from, transformation output)

The TALTIO XSLT stylesheet may be tested e.g. in an open source tool named Kernow[35].

8.2   Manual entries

There may be a need to make manual entries such as bookkeeping adjustments etc. In a production environment the user may use any existing tool with a TALTIO interface to bring manually created entries to the data warehouse.

---

[30] https://www.iso20022.org/standardsrepository/public/wqt/Description/mx/camt.053.001.02
[31] http://www.finanssiala.fi/finvoice/Sivut/default.aspx
[32] https://www.tieto.fi/palvelut/liiketoimintaprosessien-hallinta/business-information-exchange/laskuhotelli-palvelun-ohjeita-ja-kuvauksia/teappsxml-versio-272
[33] http://docs.oasis-open.org/ubl/UBL-2.1.html
[34] http://www.finanssiala.fi/finvoice/Sivut/default.aspx
[35] https://sourceforge.net/projects/kernowforsaxon/

In the POC user interface the user may add manual entries by creating correction entries. This functionality provides the means to correct some aspect or add some details to the automatically created entries (see figure 8.2.2 below). The functionality is triggered in each line clicking "edit". All entries may also be deleted by clicking the neighboring "delete". The functionality was necessary to create for the means of POC and most likely is necessary to have also in a real life system but it should not be used but only in special occasions as data should be taken in as it is in the original voucher.

Figure 8.2.2. Manual entry creation/correction in the TALTIO DW user interface

8.3   Data queries and reports from the database

The following screen captures shall demonstrate the data querying functionalities of the user interface. The logic is described in the previous chapter.

Figure 8.3.1. Filter transactions by the selected database field/-s

Source journal id	Account main id	Account sub id	Type of subaccount	Reporting code	Reporting code set identifier	Amount	Posting date
sj	1700			1700	rapko2015.1	201.54	
sj	1700			1700	rapko2015.1	201.54	
						Amount total: 403.08	

Figure 8.3.2. Listing of open receivables

Source journal id	Account main id	Account sub id	Type of subaccount	Reporting code	Reporting code set identifier	Amount	Posting date	Memo
pj	29060			29060	rapko2015.1	-59.00		active
pj	29060			29060	rapko2015.1	-74.62		active
pj	29060			29060	rapko2015.1	-429.99		active
pj	29060			29060	rapko2015.1	-435.00		active
pj	29060			29060	rapko2015.1	-100.01		active
						Amount total: -1,098.62		

Figure 8.3.3. Listing of open payables

surable code ription	Measurable id	Measurable quantity	Measurable unit of measure	Measurable cost per unit	Tax amount	Tax perce rate
uden tuloksen minen taseeseen						
uden tuloksen minen taseeseen						
uden tuloksen minen taseeseen						
					Tax amount total 145.96	

Figure 8.3.4. The amount of the period's VAT, to be paid or refunded

Add to database   Download   Search all   Filter   Open receivables   Open payables   General journal   General ledger

143 Results

Source journal id	Account main id	Account sub id	Type of subaccount	Reporting code	Reporting code set identifier	Amount	Posting date	Memo line	Authority code	Id
gi	709100			709100	rapko2015.1	51.95	2017-06-06	active		C
gi	19100			19100	rapko2015.1	372.00		active		
gi	1700			1700	rapko2015.1	-372.00	2017-06-07	active		P
gi	19100			19100	rapko2015.1	-25.44		active		
gi	29060			29060	rapko2015.1	25.44	2017-06-08	active		IS Tä
gi	19100			19100	rapko2015.1	-6.00		active		
gi	1919x			1919x	rapko2015.1	6.00	2017-06-09	active		F V.
gi	19100			19100	rapko2015.1	-190.96		active		

Figure 8.3.5. General journal

Source journal id	Account main id	Account sub id	Type of subaccount	Reporting code	Reporting code set identifier	Amount	Posting date	Memo line	Authority code	
gj	10600			10600	rapko2015.1	1585.98	2017-06-06	active		
gj	111000			111000	rapko2015.1	62256.48	2017-06-06	active		
gj	1160x0			1160x0	rapko2015.1	23461.32	2017-06-06	active		
gj	14500			14500	rapko2015.1	350.00	2017-06-06	active		
gj	1531			1531	rapko2015.1	2686.00	2017-06-06	active		
gj	1700			1700	rapko2015.1	886.52	2017-06-06	active		
gj	1761			1761	rapko2015.1	242.83	2017-06-06	active		
cr	1762			1762	rapko2015.1	140.83		active		
gj	17700			17700	rapko2015.1	303.98	2017-06-06	active		
gj	19100			19100	rapko2015.1	6650.96	2017-06-06	active		
cr	1919x			1919x	rapko2015.1	-19.44		active		

Figure 8.3.6. General ledger

8.4 Report generation with XKUBED (external tool)

The accounting entries are transferred from the database to the XKUBED tool. Alongside with the entries data the tool uses a special mapping file also presented in TALTIO format that needs to be provided as well (see figure X). The mapping tells the tool which accounts should be outputted to which element in the XBRL financial report. When the SBR code set is referenced in the accounting entries, one single mapping will serve all report preparers.

Please select your data file:

[Valitse tiedosto] **taltio_entries.xbrl**

Please select your mapping file:

[Valitse tiedosto] **taltio_mapping_v2.xml**

[CANCEL] [LÄHETÄ]

Figure 8.4.1. Uploading the TALTIO entries and mapping to XKUBED tool.

The tool shall process the uploaded instances and aggregate all entries with same account reference and output the XBRL reports that were included in the mapping file. Finally the generated XBRL reports may be downloaded and redirected to the regulatory reporting system.

Taltio

Your files from this session

File Name	Date Created (YYYY-MM-DD HH:MM:SS)	File Age	File Size (kb)
consolidatedInstance.xbrl	2017-12-05 08:06:35	0:02:55	12
dataInstance.xbrl	2017-12-05 08:06:28	0:03:02	6
FRinstance.xbrl	2017-12-05 08:06:36	0:02:55	5

Right click and 'Save link as' to save the file to your machine.

RETURN TO DASHBOARD

Figure 8.4.2. XBRL financial reports as a result of the processing in the XKUBED tool

8.5   Report submission to the central reporting gateway (Tilinpäätös 2.0)

The Finnish government gateway to receive and process XBRL data is called Tilinpäätös 2.0 and it was produced as a joint effort of the Finnish Tax Administration, Finnish Patents and Registration Office and the Government ICT Centre. The system was created in 2015-2016 and is available as a testing environment but not in production level yet. All documents produced in the demo can be found in the attachment 2 "TP_2_0" subfolder.

Figure 8.5.1. The XBRL report is uploaded to the government gateway Tilinpäätös (TP) 2.0.

The report preparer submits the created XBRL report into Tilinpäätös 2.0. (government gateway). In the case of the POC the created XBRL report essentially contains the trial balance of the financial period.

The TP 2.0. service is indifferent to the kind of financial report or the level of detail it contains. In the case of the report uploaded is the type of trial balance (extended financial report), the system enriches the data and creates the individual reports (financial reports and tax reports). The XBRL tax reports are automatically calculated and formed based on the mappings defined alongside the SBR code set.

Figure 8.5.2. Generated reports from the trial balance XBRL report uploaded to the TP 2.0 service.

The automatically generated reports include:

- tax reports ("Tuloveroilmoitus" and "Erittely varauksista, arvonmuutoksista ja kuluvan käyttöomaisuuden poistoista") and
- the financial statement ("Suppea tilinpäätös (Oy)").

The system performs XBRL validations to all of the reports and notifies the user on found issues. These may include:

- mandatory classified missing pieces of information,
- failed calculation validations (e.g. assets don't match equity and liabilities) or
- other illogicalities.

The user may provide manually additional information to the reports created in the system. Typically these include notes to the annual report, directors report, auditor's report or other textual information that may not be concluded form the trial balance. However the user is also able to submit these information in the original XBRL instance file it their own filing system is able to include these details to the report.

Figure 8.5.3. Providing additional details manually to the created reports.

The created reports are automatically groupped to draft reporting packages per report receiving regulator (or other intrest party in the future). The user may also provide additional attachment documents to each reporting package the types of which are enumerated.

Figure 8.5.4. Providing attachment documents to reporting packages (Verohallinto = reporting package for the tax reports, PRH= annual financial statement reporting package)

The user may obtain all the reports by clicking the icons 'XBRL' or 'HTML' (inline XBRL). The generated XBRL reports in the demo can be inspected in the folder: Demo materials\TP_2_0\output.

Finally the individual report packages; financial statement and tax reports are transferred to their corresponding regulators when the report preparer hits the "send" button[36]. There may be any number of regulators or other interest parties announced as report receiving entities in the service.

---

[36] Tilinpäätös 2.0 not in production mode but functional, test environment shall be used instead

9   Summary and key findings

9.1   Data quality and completeness

Real-life business data was used in the TALTIO POC. All the business documents where acquired in structural digital format, some of which had to be digitized manually. Quality inspections and corrections were also performed manually to insure the functionality of the following phases in the data handling. The following corrections were made on the source business documents:

Payment references: the same payment reference should be found in the invoice as well as the bank account statement thus enabling the reconciliation process. For an unknown reason the elements in each corresponding business document were not always containing the correct reference number or at times no data at all. There were also some occasions when the given reference number would contain forbidden characters like " ' " for instance. This is a data quality issue that derives in either the invoicing or invoice transferring system or in the banking system and therefore had to be manually corrected for the POC.

Sales invoices: in some of the source data there was a key peace of information missing which was the seller bank account ID. For the invoice recipient the payment is impossible to do without knowing to which account it should be appointed to. In these occasions this information was included in the PDF-version of the invoice but not the structural version which leads to problems in automated invoice handling.

On top of these, all account information was added manually as the invoice data was not complete in this aspect. However with one invoice it was demonstrated that if the account details are provided, they are also present in the TALTIO presentation.

In most of the business documents it's obvious that the full amount of data is not included (e.g. product information, measurable and various identifiers that could be provided). For the POC this was not perceived as an inhibitory factor for any of the functionalities to achieve in the POC but it is one for the full vision to become true.

9.2   Sign management

There are various ways on how to handle entry sign (+/-) management in accounting systems. Some include the sign of the amount to the entry amount while other present numbers as absolute values and provide additional debit/credit identifiers.

The meaning of debit and credit entries in accounting is well defined and clarified but the usage of different ways of expressing it may create confusion. For instance in the perspective of the bank account statement the debit and credit amounts are opposite in meaning to bookkeeping.

9.3   TALTIO format performance

TALTIO was perceived as a functional and expressional format definition in the scope of the POC and was performing as expected. The XSLT stylesheets had to be improved a little over the course of transforming the source data into TALTIO format as they have not overcome a full testing period yet.

The short-coming of TALTIO is mainly its current technical presentation which is XML. There might become issues with data processing capabilities when the set of entries is in a scale of a larger enterprise. However 99% of the companies are very small and their business data also remains in reasonable size. There is a new XBRL specification about to be released[37] that enables XBRL contents to be presented in other formats besides XML. These include JSON and CSV that are known to be very light on the memory consumption.

10 Conclusions and suggestions for further development

10.1 Increasing the adoption level of structural business documents

Business data may be claimed as structural given in any machine readable format. However this does not make it understandable to others. In many countries there is for instance a vast amount of different structural invoice formats used which is a hindering factor for the over all adoption as the benefits for having the structural data are harder to achieve. Standardization should be increased across domains in structural business documents to over come this.

Public sector parties have been known to declare receiving only structural electronic invoices which have had a very positive effect on the software support and adaptation of eInvoicing in general. Such actions should be continued in the future and expanded to cover also electronic receipts and orders. Where applicable, public sector actors should also set example and demand standardized interfaces. In time it should also be evaluated if a national mandate could be given on some or all of these electronic practices.

10.2 Improving data quality

In many cases there is important information missing in the structural business documents whereas in some cases important data such as the tax codes or product information are placed in to free text elements in the structural invoices. In general any receiving systems are not able to obtain the data from anywhere else but the expected designated element to hold the certain piece of information. The usage of

---

[37] https://www.xbrl.org/tag/oim/

fully tagged data should be encouraged in order to prevent the loss of data and to enable data analytics.

The Finnish law of the invoice contents is strict on details to be included in the invoice. However if the invoice is sent in both PDF- and in structural format, the requirements are often fulfilled if the PDF version but not in the structural one and in the eyes of the law this is acceptable. Such loop holes should be closed as there are such shortcomings in the data completeness of the structural invoices.

Each standard or data format definitions used in the POC provides an implementation instructions that should be followed closely by the users. Lack of surveillance seems to have led to a result where data quality on any business document received from external parties is difficult to rely on.

One solution to overcome this is to provide mechanisms to data quality assurance on the selected set of data formats. This could be for instance a national business document validator that would provide assurance on the level of data quality.

10.3 Harmonizing common data dictionaries and code sets

In order to facilitate the advanced usage of data in any level, transactional or external reporting, there should be common understanding of terms and their meaning. To achieve this goal the common data dictionaries and code sets should be harmonized and standardized.

The usage of the harmonized classifications and code sets with-in the business documents should be encouraged by the regulators. The deeper the adoption of the standards in the business processes, the more interoperable the systems will become.

10.4 Integration and automation of regulatory reporting

Administrative burden in regulatory reporting is a result of proprietary solutions defined only to serve the processes on one particular regulator. In an ideal world companies should not have to deal with additional requirements to create reports to the regulators but rather granting them an access to a view on their data that this particular regulator is entitled to see.

TALTIO together with the standard business reporting approach is already enabling this as seen by the POC so all regulators should be encouraged to work together in order to achieve the level of harmonization as the bases that SBR requires.

10.5 Standardized interfaces

For future applications there should be no pre-assumptions made on where or how the data is stored. Multiple parallel databases, a centralized data warehouse or

distributed ledger technologies may be used but even then interoperability remains key.

TALTIO doesn't take a stance on the architectural decisions but provides a functional data interface for all approaches. It should therefore be enforced to comply with the standardized interface whilst allowing the actual underlying business processes and systems to act the way they are set up.