

Roadmap for the introduction of data-based reporting 2021–2024

business analysis

Procurement reference no. 231015



MAJANDUS- JA
KOMMUNIKATSIOONI-
MINISTEERIUM



Euroopa Liit
Euroopa
Regionaalarengu Fond



Eesti
tuleviku heaks

Contents

Introduction	3
1. Definitions	5
2. Summary of the mapping of the current situation of reporting	7
3. Development proposals and general activities in the introduction of data-based reporting 10	
3.1. Creation of the legal basis for data-based reporting.....	10
3.2. Forming an overview of the data used for reporting	11
3.3. Implementation of existing solutions	13
3.4. Development of new software components by the state	13
4. Methodology for the assessment of the expenses of a state agency in the development of reporting	15
5. Management of the inter-agency taxonomy	17
6. Check-list for state agencies to support the performance of the reporting operations	17
7. Roadmap on a time scale	19
Annex 1. The agencies requiring periodic reporting and the number of reports	20
Annex 2. Legislation examined	22
Annex 3. Description of the online survey of the state agencies	23

Introduction

This document describes the action plan or roadmap of the implementation of data-based reporting. By reporting, we mean the periodic reporting submitted by businesses to state agencies, i.e. the reporting which is submitted repeatedly in time with new data and is directly or indirectly based on the legislation of Estonia.

The businesses subjected to the reporting obligation and the number of reports submitted by them (how many reports, how many data fields per report) were mapped within the framework of the work. The real data volume of the reports or the data content were not analysed.

In the course of the analysis, it was identified that Estonia has already been moving towards alleviating the administrative burden of businesses and towards data-based reporting for years (e.g. within the framework of the 'Aruandlus 3.0' project). The most important problem to be solved is currently an infrastructure that would meet the needs of businesses. Statistics Estonia has developed an infrastructure consisting of the data reception service and guidelines for some datasets, but there is no feedback mechanism which would allow an entrepreneur to check whether or not they have submitted data appropriately. This is a significant shortcoming, as **it is not possible to develop the interfaces due to the lack of feedback – it is very difficult to test the system if the recipient of the data does not return information about the deficiencies of the data submitted**. Thus, organising data reception in a manner which would meet the needs of the businesses is the most important problem for the state which prevents the implementation of data-based reporting pursuant to the rules of Aruandlus 3.0 by the businesses. In the reception of data, all state agencies should use a uniform technological solution compliant with the needs. This would ensure a significant effect in alleviating the administration burden. Thus, the Ministry of Economic Affairs and Communications should undertake operations with all state agencies all at once to introduce data-based reporting and reduce the administrative burden, in order to implement a uniform reporting infrastructure at all agencies. Then, when the infrastructure is in place, the economic impact of adding each specific report to the new infrastructure should be assessed as the next step, the report should be organised based on priority on the basis of the analysis, and transfer of the reports based on priority should be performed.

This document was drafted based on the vision of real-time economy¹ and the mapping of the reporting obligation created within the framework of the project commissioned by the Ministry of Economic Affairs and Communications.

Real-time economy is a digital ecosystem in which the transactions between different parties occur in real time or with a minimum delay. This means replacement of paper-based economic transactions and administrative operations with automatic data exchange in a digital, structured, machine-processable, and standardised format.

¹ https://www.mkm.ee/sites/default/files/reaalajamajanduse_vision_2020-2027_kaskkiri.pdf

Based on the vision, we only discuss the reporting which is based on structured data in the document. Another significant restriction is discussing the periodic reporting arising from the legislation of Estonia. The document does not cover short-term reporting between certain parties which arises from the performance of a contract. The reporting with a wider circle of related parties is discussed, the atomisation of which and the development of a machine interface may be economically reasonable, i.e. it would allow saving the resources of the state and the businesses. In order to facilitate inter-agency reuse of data, creating data descriptions for individual reports may also create value in the future (for example, in the next public procurement) if such reports are used by many businesses.

Based on the technical specifications of the public procurement, the roadmap must meet the following requirements:

- 1) The development and action plan required for the creation and standardisation of the uniform taxonomy of the data compositions which the reporting is based on (incl. for the implementation of XBRL GL) was drawn up. The roadmap is a planning tool which includes a list of actions with a schedule and the methodology for planning the volume of the works for the transfer to a uniform inter-agency taxonomy.
- 2) The roadmap provides to the state agency a more efficient sequence of operations for performing the works and a methodology for the assessment of a potential need for resources. The roadmap includes development proposals for creating the capabilities for the reception and processing of machine-readable reporting data and for the reuse of the data. The roadmap also includes suggestions for methods for the management of an inter-agency taxonomy.
- 3) The roadmap helps state agencies to save on excessive repeating of operations and to maintain a uniform inter-agency focus in the developments.

The project is co-funded from the European Regional Development Fund.

1. Definitions

Definition	Explanation
Data processing	Creation, reading, amendment, deletion, transforming, or exchange of data (organising of data traffic).
Classification	<p>A classification is a list of unique codes and names designed for the classification of phenomena. Each element of a classification must have a code and a name and an element may also have further attributes, such as names in other languages or a feature that makes the classification of the element more easily recognisable. Classifications may be linear or hierarchical. In a linear classification, there are no connections between the elements of the classification. In a hierarchic classification, the elements may in turn be divided into elements, i.e. there may be hierarchic relationships between the elements.</p> <p>In this document, classifications include both national classifications as well as other lists, incl. code lists designed for the classification of data objects nationwide or within one specific report. Thus, the concept of a classification is wider in this document compared to other situations and the law on the system of classifications</p>
Classification element	A unique sub-part of a classification defined by the code name and other data attributes. The codes of different classification elements do not overlap.
Process	A sequence of operations. A process is divided into operations which may be in the linear as well as recursive or branching sequence. Recursions may cause repeated performance of the same operation. A sequence of operations performed based on pre-determined rules can be referred to as a process. Any sequence of unrelated operations is not referred to as a process in this document.
Schematron	<i>Schematron</i> ² is a rule-based data validation language which enables identification of the presence or lack of patterns in XML data. <i>Schematroni</i> is used for defining and implementation of the inspections of XML data to determine the quality of a specific XML file based on the rules defined.
Taxonomy	In this context, taxonomy is the description of a dataset which includes the list and descriptions of the data objects and data elements, as well as classifiers and other important information about the structure and content of the dataset. Taxonomy or the content of the description of a report is described in detail in

² <https://schematron.com/>

Definition	Explanation
	the following section of the reporting model document: '2.2.2. Outcomes of describing the composition of data'.
X-tee	X-tee is a technical and organisational environment which allows for secure online data exchange between state agencies and with the public sector which ensures evidential value. ³
XBRL GL	XBRL GL or <i>XBRL Global Ledger</i> is special structure of XML which is oriented to the submission of reporting data. While XBRL is mainly oriented to saving aggregated data, XBRL GL also enables saving individual records, for example the data of economic transactions by individual transactions. XBRL GL enables the submission of data at the level of individual records. It is important to point out that XBRL GL is the most efficient for the real-time submission of data and for cases in which the number of individual records to be sent together is not very high.
XML	XML ⁴ is a tool for saving and transporting data and it is independent from hardware and software.
Individual records	A data entry about a specific subject or object. An individual data entry does not include summarised records.

³ <https://www.ria.ee/et/riigi-infosusteem/andmevahetuskiht-x-tee.html>

⁴ https://www.w3schools.com/xml/xml_what.asp

2. Summary of the mapping of the current situation of reporting

In course of mapping the situation of reporting, the reports which are requested **repeatedly** by agencies **from the businesses within an unlimited period of time** were considered. The submitters of the data are either all Estonian businesses or those businesses which meet the criteria described in a specific piece of legislation. The unlimited period of time means that the reporting obligation has been established by the piece of legislation for an unspecified term. Any reporting related to specific events was not discussed in detail. For example, whether or not the receipt and use of a benefit, a shipping accident, environmental pollution, etc. include reporting. The later were not discussed, as the event or phenomenon which gives rise to the reporting obligation is of a non-permanent nature. This, in turn, means that it would not be feasible for the reporting party to create a separate system for such reporting, as the volume of the development work would exceed the volume of the work required for drawing up an individual report. Thus, the reporting will not be automatised from the perspective of a business, as this would not be economically feasible. A state agency may, however, have created a respective information system used for accepting and processing event-based data.

The situation of the reporting was mapped based on online surveying of state agencies and an analysis of legislation. The information obtained from the online survey was specified based on the content of legal acts and regulations. The list of the pieces of legislation reviewed is provided in [annex 2](#) to the document. The questionnaire of the online survey can be found in [annex 3](#) to the document. The details of the reports identified based on the survey and the analysis are provided in [annex 1](#) to the document.

An interactive overview of the reporting is available in the [online application of the mapping of the reporting](#)⁵. The online application can be used to filter the mapping by different dimensions and to obtain report-based or state agency-based results.

Next, we would like to highlight the main facts about the reporting identified in the course of the analysis.

Thirty-one agencies, in total, require periodic reporting. The data requested are divided into 421 reports. The reports include about 60,000 reports, in total. The reporting arising from the directly applicable EU directives which the businesses must also submit is added to these reports. A detailed overview is provided in [annex 1](#) to the document. This result was obtained based on an online survey organised at state agencies and based on an analysis of the **Estonian legislation** concerned with reporting.

The highest number of sets of data is collected by Statistics Estonia: 111 datasets, in total. Statistics Estonia is followed by the Financial Supervision Authority with seventy-nine reports based on the Estonian legislation (and a further 256 reports based on the EU directives directly applicable in Estonia, which were not analysed further in the project) and the Bank of Estonia with thirty-eight reports. The remaining agencies collect considerably less reports. This result was obtained by questioning forty-two state agencies and analysing 793 pieces of legislation.

⁵ https://tietoanalytics.ee/MKM_aruanded

The X-tee can currently be used by the businesses to submit the data specified in Table 1.

Table 1. The data transmission capacity created for the X-tee

Agency	Report name	The channel/data exchange platform used for submitting reports
Tax and Customs Board	Gambling tax declaration for organising a game of change as a ring game	X-tee, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	Gambling tax declaration for organising a toto	X-tee, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	Gambling tax declaration for organising a game of chance or a game of skill in the form of remote gambling	X-tee, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	Gambling tax declaration for organising a game of chance on a gaming machine	X-tee, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	Gambling tax declaration for organising a game of chance on a gaming table	X-tee, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	Gambling tax declaration for organising a tournament of game of change	X-tee, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	IOSS – the declaration for E-commerce and services (the import procedure)	X-tee, uploading a file, self-service environment
Tax and Customs Board	Value added tax return	X-tee, uploading a file, self-service environment, e-mail, regular mail, on-site
Tax and Customs Board	Warehouse entries of temporary storage premises for storing fuel, a customs warehouse, free zone, inward processing, excise warehouse, tax warehouse, provider of storage services	X-tee
Tax and Customs Board	OSS – the declaration for E-commerce and services	X-tee, uploading a file, self-service environment

Ministry of Education and Research	Various data in the Estonian Education Information System	X-tee, uploading a file, self-service environment
Tax and Customs Board	Tax returns for the income and social tax, contributions to mandatory funded pension, and unemployment insurance premiums	X-tee, uploading a file, self-service environment, regular mail, other (from an online bank over X-tee upon the payment of wages)

In the course of mapping the reporting, the hypothesis of the needs of the reporting and the data being described in very different formats by the agencies or not being described at all was confirmed. Machine-interfaced reporting which would be even more based on individual records is basically not used in Estonia today. Neither the private sector nor the state agencies are capable of data-based reporting. Statistics Estonia is an exception, having made the first developments in this direction. Another exception is the Financial Supervision Authority, which accepts the data of all transactions in shares executed in Estonia.

Businesses generally obtain the data for drawing up reports in a semi-automated form, but private businesses compile the data submitted from different systems and data sources and, as a rule, send the data via the self-service portal of the state agency. Automatic submission of reports is above all too expensive from the perspective of the businesses submitting the data, as it calls for the development of various different systems by the business. Automatic transmission and the idea of data-based reporting is accepted better by those software and accounting service providers who have either created the required software themselves or are using systems with better options for adjusting the outputs and interfaces. Submission of the economic data of commercial undertakings is currently the only field in which automatic data transfers have been introduced.

Thus, for the data-based reporting over X-tee to gain ground in the private sector, state agencies should make the submission of data over X-tee very convenient and similar for different agencies. The state should create and implement the software components described in the roadmap. Businesses are used to self-service environments and changing this habit may take time, as development of the systems is much less wide-spread in the private sector compared to state agencies. Those software service providers, incl. Enterprise Resource Planning (hereinafter ERP), which are able to generate an economic effect with the help of a higher number of data submitters and have tested data-based reporting are an exception. In the case of ERP, there are more businesses together behind one development and one X-tee connection, which helps to save on the expenses in the wider perspective. Providers of ERP services have not yet reached further from the testing stage with data-based data exchange, as the state has not completed the development of a solution for receiving the data – there is no data feedback mechanism.

3. Development proposals and general activities in the introduction of data-based reporting

A working group should be formed for the management of the introduction of data-based reporting and for consultation. The working group should find solutions to the following issues:

- 1) Which agency will manage and develop the intra-agency taxonomy?
- 2) Which agency will manage the nationwide set of descriptions of the reports?
- 3) Where (in which system) will the taxonomies of the data be stored?
- 4) Who will be responsible and how will the taxonomies be updated to burden the businesses as little as possible?
- 5) Where will the guidelines of data-based reporting be published and kept?
- 6) How will the classification be managed and the data made available? Centrally (and at which agency) or separately at each agency?
- 7) Which agency will procure and provide to other agencies the software components listed in the roadmap?
- 8) How will the legal framework for the data-based reporting be created?
- 9) What kind of a legal framework (and legal basis) is required for the creation and management of the data descriptions of the reports?

Any issues arising in the course of operation will be added to those listed above.

3.1. Creation of the legal basis for data-based reporting

The need to amend the legislation is one of the most significant development needs in the implementation of individual records based or data-based reporting. Individual records include personalised data. These may enable identifying persons or their business secrets. If the state wishes to access any personalised datasets which include personal data as well as a business secret, this should be legalised by legislation. Today, Statistics Estonia, which has the mechanisms required for data protection, may use individual records. Statistics Estonia is prohibited by law to disclose data to other agencies. Individual records are also used in keeping different registers.

In the currently used reporting in general, leaving aside some reports of Statistics Estonia and the Financial Supervision Authority, individual records are not used and there are no legal grounds for the collection thereof. These grounds should be established if the collection of individual records is justified and useful for the society as a whole from the economic or social perspective.

One option for the legalisation of the use of individual records by the state is to create a respective framework act and specify the rules by regulations in each specific area of administration depending on the specifics of each dataset. A separate legal analysis should be conducted to find a specific solution.

There is currently no legal basis for the creation and systematisation of data descriptions for reports. Agencies are currently not obligated to create and publish data descriptions for reports, even though this is necessary to ensure the data quality of the reports and to reduce the workload of the agencies themselves.

3.2. Forming an overview of the data used for reporting

The online survey showed that many state agencies are not capable of immediately responding to the question of which data and reporting they are using today. Getting an overview of the reports is especially difficult in the case of those agencies which have recently been formed by merging different agencies (e.g. the Education and Youth Board) or in which the IT services have been recently consolidated.

For the reuse of data to gain ground, it is first necessary to organise the system for getting an overview of the datasets of the reports, i.e. the taxonomy catalogue in Estonia. The necessity of the data descriptions has been discussed at the state level before, primarily in the light of documenting the datasets of databases. The information on which datasets private undertakings are required to collect periodically is not currently systematised at the nationwide level. Statistics Estonia has described which variables are collected, but most other agencies do not yet have such clear systems. Statistics Estonia is mainly involved with the documentation of the data which are required for compiling nationwide statistics.

Thus, for data-based reporting to function in an optimised form, all descriptions of the data used in reporting, i.e. the descriptions of the data compositions must be publicly disclosed in a systematised manner. This means that the dataset collected must be documented in detail and each data element must be searchable by content and name. The search function is required to prevent the need for reporting. If the data consumer can find the data element required from the taxonomy catalogue, they can request the data from another agency without burdening the businesses.

In order to prevent double description, the structure of the taxonomy catalogue must be aligned with the structure of the descriptions of the data compositions at the level of detail (see the reporting model in section 2.2.2).

In the documentation of reporting data, it is important to make sure that not the national datasets, but the input data needed for forming the datasets are documented. Thus, these data are also not directly included in the dataset described in the administration system of the state information system (hereinafter the RIHA)⁶. The data descriptions in the RIHA are not at the level of a variable or a data element, but at the level of data objects. Thus, information is only available from the RIHA at a very general level (see

⁶ <https://www.riha.ee/>

examples from the 'Andmed' ('Data') information system description block in the RIHA⁷) and only if the description of the dataset has been updated. The current requirements for describing data in the RIHA are established by the so-called RIHA regulation⁸. Specifications of the requirements are provided in the instructions⁹. There are plans to implement the new requirements by the guidelines drawn up by Statistics Estonia approved on the basis of an amendment to the Principles for Managing Services and Governing Information¹⁰. It is now possible to enter detailed descriptions of data in the RIHA, but different datasets have approached this differently. Many dataset managers have uploaded detailed descriptions into the RIHA.

In addition to national datasets, there are also datasets managed by private operators for fulfilling public duties and the data compositions of those datasets have also been described in some form. Accounting businesses have also done great work in the field of standardisation, having defined the data composition of the accounting records in 2020¹¹ and connecting it to XBRL GL. **The description of the data composition of the accounting records should be included in the composition of the taxonomy catalogue of accounting data in an upgraded format.** It would be usable as the data composition of several current national reports (such as the KMD or TSD tax declaration forms) at the level of individual records.

The existence of the taxonomy catalogue of reporting data or the data collected from private businesses helps to significantly reduce the need for business analyses and data analyses which are conducted for involving new data in the work of an agency.

The taxonomy catalogue must be connected to the application receiving data so that if the data structure is changed in the application, the taxonomy is also changed automatically. When adding data, the descriptions must also be added manually. If the application for receiving data is not connected automatically, the application and the taxonomy catalogue may go out of sync with the first changing of the data composition, in which case the catalogue is of little use. The existence of data descriptions must be checked automatically so that the application cannot function if the data element does not have a proper description.

It is reasonable to organise the management of the taxonomy catalogue centrally with one agency managing the data descriptions and ensuring that they are updated, but delegating the development of the data compositions to the agencies which need the dataset. The central manager will provide instructions and check describing of the data.

Management of a taxonomy catalogue functions best if the updating of the data descriptions of the taxonomy catalogue is automated. Updating of the data structures from the real database or from real XBRL GL data must be automated. In this case, the changes in the data compositions made in the data

⁷ <https://www.riha.ee/Infos%C3%BCsteemid/Vaata/pais#andmed>

⁸ <https://www.riigiteataja.ee/akt/129032016006?leiaKehtiv#para18>

⁹ <https://abi.ria.ee/riha/kirjeldamine/andmed>

¹⁰ [Principles for Managing Services and Governing Information – Official Journal](#)

¹¹ <https://wp.itl.ee/wp-content/uploads/2021/02/ANNEX-I-XBRL-GL-accounting-entry-data-standard.pdf>

exchange in the course of development will not remain unnoticed. A verbal explanation of what the added data element entails must be added manually. It is difficult to execute automatic updating of the data composition based on the XBRL GL real data, as the data submitter may enclose data elements which the agency has not requested with the data. Such elements cannot be treated as changes in the data composition. It would be more reliable to scan the changes in the data composition from the database of the stage agency where the treated data finally arrive. This also the case if the database is only designed for storing reporting data. If data are transferred into different databases and mixed with other data, automatic scanning for changes is no longer feasible from the perspective of the data descriptions of the report, as the changes may concern data which are not related to the reporting.

3.3. Implementation of existing solutions

In order to prevent needless expenses, the following existing solutions should be taken into use:

- 1) **The X-tee infrastructure**¹², which is currently used little by private undertakings (see the statistics on the use of X-tee ¹³), as it is relatively expensive and not very easy to use from the technical perspective.
- 2) **Statistics Estonia has created a software**¹⁴ **of the X-tee service for receiving data**¹⁵, which is an open-code software and should be implemented by all state agencies to make the interfacing of data reception identical in the case of all agencies from the perspective of a private undertaking.
- 3) **Statistics Estonia has created REST services of classifications**¹⁶. Statistics Estonia is using a software for issuing information of classifications which should also be taken into use by other agencies. If all state agencies start using the same software, businesses can use the same interface for all state agencies and the investment required for interfacing with each individual agency is considerably smaller.

3.4. Development of new software components by the state

In addition to the universal data reception service developed by Statistics Estonia, the following should also be developed:

- 1) **A module for checking the data submitted** which would implement the checks created in the course of describing the data composition of the data in the *schematron* (etc.) format and save the database data quality information which could later be issued to businesses.

¹² <https://www.ria.ee/et/riigi-infosusteem/andmevahetuskiht-x-tee.html>

¹³ <https://x-tee.ee/stats/>

¹⁴ <https://www.riha.ee/Infos%C3%BCsteemid/Vaata/estat>

¹⁵ <https://koodivaramu.eesti.ee/statistikaamet/xgate-client>

¹⁶ <https://estat.stat.ee/codelists/availablecodelists/>

- 2) **An X-tee service for reflecting the data quality** which would provide the system of the business information about the outcomes of checking the results.
- 3) **An X-tee service for reflecting changes in the content of the data** which would give the data submitted feedback on changes made in the data by the state agency. The use of this service is important if the data submitted come with obligations for the business, such as a tax liability. If the Tax and Customs Board changes the amount of the tax based on the data submitted, the business must be able to learn about this automatically without entering the portal of the state agency.
- 4) **A solution for entering data in the system of a state agency (the ETL¹⁷ tool)**. This part of the system is required for the data which a state agency has received to arrive in the required place in the database of the agency. It is convenient to transport data in the XBRL GL format, but the data are more readily available for information systems in a database. The acquisition of a uniform nationwide basic tool should be considered here. In the event of a choosing a standard ETL tool, the expenses on the configuration of the tool in the environments of the specific agency should also be taken into consideration. In the event of more complicated and voluminous and real-time data, however, the agency may need to acquire a specific tool for their processes and systems which is developed based on the requirements of the agency.

¹⁷ https://en.wikipedia.org/wiki/Extract,_transform,_load

4. Methodology for the assessment of the expenses of a state agency in the development of reporting

The following methodology applies provided that a standard reporting solution is used which is described in this document, i.e. the recommended existing software components have been taken into use and the required additional components have been developed and introduced. In this case, the numbers of reports and data fields can be used as the basis measures for the assessment of the expenses of the state agency. Those basic measures can be used if the agency has a general idea of which data are required and the data needs have been documented at the general level or in the form of a 'paper-based form'.

An external development and implementation process mainly consists of the following components:

- 1) system analysis;
- 2) development and testing;
- 3) general operations, such as project management, testing, introduction, maintenance, etc.

In the course of system analysis, the data must be modelled in detail based on the reporting model; this includes detailed description of the XBRL GI-based structure of the dataset, the business rules/data checks, and other semantics specified in the model. The reflecting mechanisms, i.e. the data checks, are executed in the course of the development. The same applies to the settings of transferring data into the system of the agency. This process does not involve examining the general software (e.g. the XBRL GL reception module) which is designed for using by all agencies.

If the agency orders the development service externally, the following formula may be used for calculating the cost:

Financial cost of the development = (number of reports * number of data fields * 1.5 + number of reports * 160) * 1.2 * average hourly fee.

For example: If the agency needs to organise the reception of three reports so that each report contains approx. 55 data fields, the cost of the development related to these reports is found as follows:

$$(3 * 55 * 1.5 + 3 * 160) * 1.2 * 60 = 52,380 \text{ euros.}$$

This formula is based on the assumption that working on each data element in the analysis stage takes 1.5 h, on average, the development of each report takes 160 h, and the general costs add a further 20% to the cost of the works.

When using this methodology, the level of preparedness of the infrastructure of the agency for the introduction of the reports should be considered. **If the systems have not yet been properly tested and used, the amounts found should be further multiplied by a higher number** to form a risk buffer for

works which may be added. The time spent by the employees of the agency which may be equal to the working time of the development partner must also be taken into consideration.

In addition to the development costs, the further cost on the upkeep of the system which largely depend on the functionality and architecture of the information system of the agency must also be considered. Management includes repairing the system, fixing errors, and data management (incl. management of classifications, complementation and improvement of the data checks and business rules applied to the data, management of the data of the reports received, fixing any issues in the functioning of the system, etc.). The extent and cost of the maintenance operations may differ by systems and the functionalities of the systems are different.

5. Management of the inter-agency taxonomy

The description of the reporting model includes the need to create a taxonomy catalogue. Management of the inter-agency taxonomy is one part of the management of the taxonomy catalogue. **The inter-agency taxonomy is the common part of all taxonomies. The agencies** collecting the same data **must communicate with one another through the taxonomy catalogue for standardisation of the reporting.** This means that the agencies must proceed from the data descriptions created by other agencies and use the sub-parts thereof in their own reports. For example, the data structures of persons, addresses, and other parts of the report must be used in a uniform manner. The classifications which have already been created for classifying certain phenomena must be reused instead of creating new classifications for each report. If the data description created cannot be used in an unchanged form, an inter-agency working group must be formed which will develop a new version of the data description which would be suitable for all users. If no consensus is reached, separate data descriptions must be created.

The inter-agency taxonomy is formed by the common part of all data compositions of the taxonomy catalogue. The common part includes the following data:

- 1) the identifiers, names, contact details of the data submitter;
- 2) the identifiers, names, contact details of the recipient of the data;
- 3) the period of the data;
- 4) the identifiers of the systems of origin and destination;
- 5) the languages and countries;
- 6) the amounts and units of measure;
- 7) the details of additional documents.

6. Check-list for state agencies to support the performance of the reporting operations

Upon introducing new reporting, the following questions should be answered:

- 1) Is the collection and use of the dataset planned economically feasible? *If the answer is 'NO', collection of the data should be foregone.*
- 2) Is asking the business to provide the data the only way for obtaining the data? *If the answer is 'NO', additional data sources found in the taxonomy catalogue described in the Model should be analysed in detail or a data audit of the source should be conducted if the dataset has not been described.*

- 3) Is there a legal basis for using the data? *If it appears that the answer is 'NO', a legal analysis of the description provided in the Model should be conducted to develop the legal bases.*
- 4) Are the taxonomy, incl. the data composition described, based on the reporting model? *If the answer is 'NO', describe the data composition based on the process and outcome and describing the data composition specified in the Model.*
- 5) Does the agency have an X-tee service for receiving the data? *If the answer is 'NO', the data reception X-tee service described in the Roadmap should be taken into use.*
- 6) Does the agency have an X-tee service for reflecting the quality of the data? *If the answer is 'NO', the developed service described in the Roadmap should be taken into use.*
- 7) Are the classifications used in the reporting managed at the agency and made available to the data submitted in a machine-readable format via REST-API? *If the answer is 'NO', Statistics Estonia should be consulted and a respective software recommended should be taken into use.*
- 8) Does the information system of the agency have the capability for using data in the XBRL GL format? *If the answer is 'NO', the respective capability should be developed in the information system of the agency or a centrally acquired ETL tool should be taken into use which enables transfer of the XBRL GL data into systems.*
- 9) Does the agency have a budget for the introduction of the capabilities discussed above if they have not yet been introduced? *If the answer is 'NO', it would be advisable to seek for potential funding opportunities in cooperation with the Ministry of Economic Affairs and Communications.*

7. Roadmap on a time scale

The following table and diagram present a general action plan for data-based reporting. The action plan is divided in two:

- 1) preparation operations;
- 2) analysis of the data compositions and implementation of XBRL GL.

Without preparation operations, there is no point in analysing the data compositions or implementing XBRL GL. There are currently several businesses among the EPR software service providers who would be prepared to use XBRL GL-based data exchange, but the state has not yet created the respective X-tee services. Only Statistics Estonia has a data reception service and the services for issuing the details of classifications. There is, however, no data feedback service, which is a great problem for businesses, hindering development and introduction of XBRL GL in the production environment.

If the preparation operations were completed, there would already be potential businesses which would be prepared to transfer their data.

After the completion of preparation operations, the existing XBRL GL-based datasets which Statistics Estonia has already been working on for years should be introduced. Introduction of the fuel and packaging datasets should also be moved forward with as soon as possible. The datasets (taxonomies) which have already been analysed can be found at <https://www.stat.ee/et/aruandlus-30>.

The operations presented in the following GANTT diagram are described in detail in the document above and in our reporting model document¹⁸.

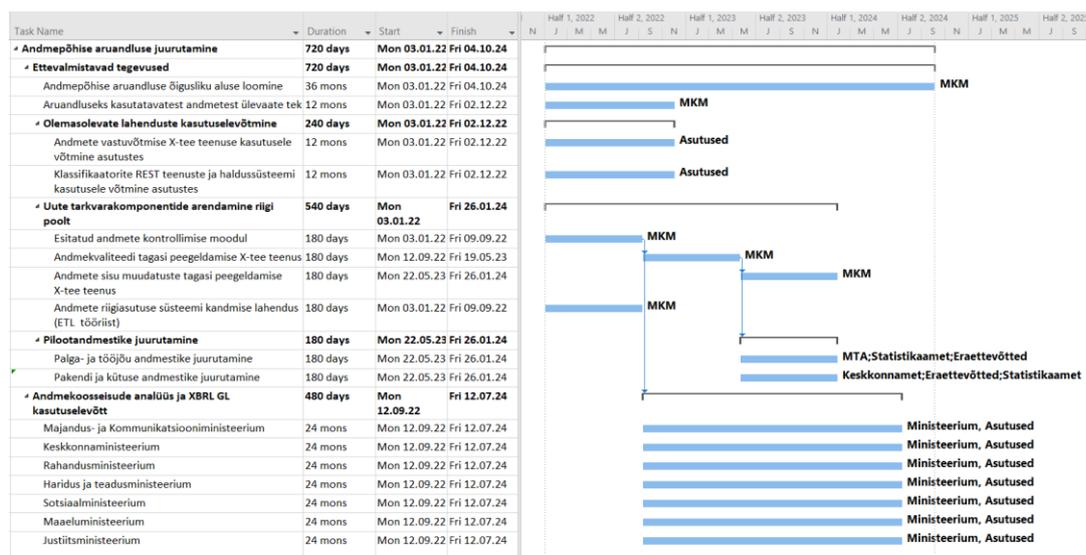


Figure 1. The timeline for the introduction of the new reporting model

¹⁸ Aruandluse mudel v1.1.docx

Annex 1. The agencies requiring periodic reporting and the number of reports

The following table includes the data as at 12 August 2021.

Table 2. Number of reports required by agencies

Agency	Number of reports
Statistics Estonia	111
Financial Supervision Authority	79
Bank of Estonia	38
Tax and Customs Board	35
Consumer Protection and Technical Regulatory Authority	28
Ministry of the Environment	22
National Institute for Health Development.	13
Health Board	13
State Agency of Medicines	11
Ministry of Justice	9
Ministry of Social Affairs	8
The Agriculture and Food Board	7
Environmental Board	6
Centre of Registers and Information Systems	6
Ministry of Education and Research	4
Transport Administration	4
Association of Auditors	3
Competition Authority	3
Ministry of Economic Affairs and Communications	3
Ministry of Finance	3
Health Insurance Fund	2
Ministry of Culture	2
Rescue Board	2
Transplantation Centre (Tartu University Hospital Foundation)	2

Agency	Number of reports
Court	1
Land Board	1
Ministry of Rural Affairs	1
Estonian Private Forest Centre	1
Ministry of the Interior	1
Social Insurance Board	1
Ministry of Foreign Affairs	1

The list of reports and the descriptions can be found in the file 'LISA 1 Leitud aruanded v1.3.xlsx' enclosed to the document.

Annex 2. Legislation examined

The table of the pieces of legislation can be found in the file 'LISA 2 Läbi vaadatud õigusaktid.xlsx' enclosed to the document.

Annex 3. Description of the online survey of the state agencies

Reporting questionnaire

The purpose of this questionnaire is to support the simplification and standardisation of regular reporting. The questionnaire is designed for the state agencies which receive periodic reports from private businesses.

The surveying is organised based on public procurement no. 231015 organised by the Ministry of Economic Affairs and Communications. The contracting authority has authorised Tieto Estonia AS to serve as the processor of the data collected within the framework of the contract.

The questionnaire is concerned with the periodic reporting collected by your agency. Please provide information about all periodic reports.

Responding to the questionnaire takes at least 30 minutes, on average, depending on the level of complexity of the reporting of the agency.

Please address any questions about the questionnaire or responding to the questionnaire to kysitlus@tietoanalytics.ee or call +372 5903 9494.

General Information

All agencies may provide information about their reports in one or several responses. It is not necessary to provide information about all reports together, the online questionnaire may be submitted several times, for example, separately for each area of activity, and by different contact persons. The details of 1 or up to n reports can be provided by submitting one questionnaire.

- 1) What agency are you responding for? *
- a. Data Protection Inspectorate
- b. Education and Youth Board
- c. Ministry of Education and Research
- d. Emergency Response Centre
- e. Ministry of Justice
- f. Ministry of Defence
- g. Estonian Internal Security Service
- h. Defence Resources Board
- i. Defence Forces

- j. Language Inspectorate
- k. Environmental Board
- l. Ministry of the Environment
- m. Competition Authority
- n. Ministry of Culture
- o. Land Board
- p. Ministry of Rural Affairs
- q. Ministry of Economic Affairs and Communications
- r. Tax and Customs Board
- s. National Heritage Board
- t. Estonian Patent Office
- u. Police and Border Guard Board
- v. Prosecutor's Office
- w. The Agriculture and Food Board
- x. Agricultural Registers and Information Board
- y. Rescue Board
- z. Ministry of Finance
- aa. National Archives
- bb. State Agency of Medicines
- cc. Centre of Registers and Information Systems
- dd. Information System Authority
- ee. Government Office - Government of the Republic
- ff. Ministry of the Interior
- gg. Social Insurance Board

- hh. Ministry of Social Affairs
- ii. Statistics Estonia
- jj. Consumer Protection and Technical Regulatory Authority
- kk. Health Board
- ll. Transport Administration
- mm. Labour Inspectorate
- nn. Estonian Foreign Intelligence Service
- oo. Ministry of Foreign Affairs
- pp. Other

2. Contact person name: *
3. Contact person's e-mail address: *
4. Contact person's phone number:
5. Does your agency establish other fixed-term obligations to businesses in addition to periodic reporting?*

By periodic reporting, we mean the reports which are submitted to the agency based on new data repeatedly at a certain interval. Single reports submitted once are non-periodic reports. Fixed-term obligations are, for example, financial obligations or obligations to perform other operations (e.g. an obligation to make a payment, perform control operations, assemble meetings). An obligation may be applicable to all or to some businesses (e.g. to the businesses of a certain sector, to the businesses which have used a certain service).

- a. Yes
- b. No

6. If you responded 'Yes' to the previous question, please specify briefly which other fixed-term obligations are placed on businesses by your agency?

For example, single reports submitted on one occasion, financial obligations or obligations to perform other operations (e.g. an obligation to make a payment, perform control operations, assemble meetings). An obligation may be applicable to all or to some companies (e.g. to the companies of a certain sector, to the companies which have used a certain service).

7. How (through which channels) does your agency provide information about the reporting to businesses? *

Which channel (e-mail, self-service environment, website, etc.) is used to notify the business of the need to submit a report, to remind of the obligation, and to notify that the obligation has been fulfilled. If the notification process differs by reports or some of the operations are not performed (e.g. the businesses which have submitted their reports are not notified that their obligation has been fulfilled), please describe this briefly.

8. Which technological solutions are used by your agency for the reporting-related operations? *

Are you using a technological solution to send notifications about the need to submit reports to the business, to monitor submission of the reports, send reminders, notify that the obligation has been fulfilled? If different technological solutions are used for different reports or if the technical solution is only used for some operations (e.g. the initial notification is sent automatically, but the submission of the reports must be monitored manually), please describe this briefly.

9. Is the agency using periodic reporting? *

Periodic reporting means the reports which are submitted to the agency based on new data repeatedly at a certain interval. If you respond 'No' and submit the form, you have completed responding to the questionnaire.

- a. Yes
- b. No

-
10. Has the agency used XBRL GL as the reporting data format? *

- a. Yes
- b. No

-
11. Please describe the XBRL GL experience in further detail. *

Please also provide examples of previous use of the XBRL GL standard. (Please do not confuse it with XBRL)

-
12. Has the agency used the Aruandlus 3.0 principle in reporting? *

- a. Yes
- b. No

13. Where and how has the agency come in contact with Aruandlus 3.0? *

14. Next, please provide some information about the reports used by the agency.

Would you like to submit the description of the reports as a file (URL) or enter the description of the reports under the responses to this questionnaire? *

Please use the Excel format to submit the data in the form of a file. The form can be downloaded here: https://tietoanalytics.ee/MKM_aruanded/aruanded.xlsx

If you would like to e-mail the descriptions of the report, please sent the file(s) to: kysitlus@tietoanalytics.ee .

- a. I am submitting the description of several reports as a file (the Excel form and the e-mail address are provided in the description of the question)
- b. I am submitting the description of several reports as a link to a website (the link can be added under the following question)
- c. I would like to describe all reports under the responses to this questionnaire.

Information about the reports used.

Below, you can submit information about the reports used by your agency as a link to a website. Please only provide information about the reports which are submitted periodically (not single reports submitted on one occasion).

Please use the Excel format to submit the data in the form of a file. The form can be downloaded here: https://tietoanalytics.ee/MKM_aruanded/aruanded.xlsx

15. Please enter the link to the description of the reports: *

Information about the reports used at the agency.

Below, you can submit the information about one report used at the agency. Please only provide information about the reports which are submitted periodically (not single reports submitted on one occasion). The descriptions of the reports must be submitted separately for each report. Having filled in the questionnaire about one report, you can start submitting the information of another report without having to fill in the general information block again.

16. What is the name of the report? *

17. How often are the businesses required to submit the report? *

- a. regularly
- b. regularly updated
- c. once per hour
- d. once in two hours
- e. once in three hours
- f. twice a day
- g. once a day
- h. once a week
- i. three times a week
- j. twice a week
- k. three times a month
- l. twice a month
- m. once a month
- n. once in two months
- o. once a quarter
- p. three times a year
- q. once in six months
- r. once a year
- s. once in two years

- t. once in three years
- u. once in four years
- v. once in five years
- w. once in ten years
- x. once in twenty years
- y. once in thirty years
- z. irregularly
- aa. unknown
- bb. never
- cc. Other

18. Which period the report is submitted for? *

- a. regularly
- b. 1 hours
- c. 3 hours
- d. 12 hours
- e. 1 days
- f. 2 days
- g. 3 days
- h. 1 week
- i. 10 days
- j. 2 week
- k. 15 days
- l. 1 months
- m. 2 months

- n. 1 quarter
- o. 4 months
- p. 6 months
- q. 1 years
- r. 2 years
- s. 3 years
- t. 4 years
- u. 5 years
- v. 10 years
- w. 20 years
- x. 30 years
- y. irregularly
- z. unknown
- aa. never
- bb. Other...

19. What is the term for submitting the report? *

E.g. Which date, which month

20. Please provide a brief description of the content of the report. *

Please describe briefly the subject and purpose of submitting the report

21. In which data medium/format can the report be submitted? *

Please specify all options available.

- a. csv
- b. doc
- c. paper

- d. pdf
- e. xbrl-gl
- f. xls(x)
- g. xml
- h. Other

22. Which channel and/or data exchange platform can be used to submit the report? *

Please specify all options available.

- a. X-tee
- b. Other machine interface
- c. Uploading a file
- d. Self-service environment
- e. Email
- f. Regular mail
- g. On-site
- h. Other

23. What is the total number of data fields which must be submitted by the reporting parties? *

The total number of data fields means the number of questions which the reporting party must respond to. If a question includes several sections, it may be counted as one data field (e.g. the amount of fuel consumed in tons, by types of fuel).

24. Which data (data fields) are requested by the report? *

Here, you may provide a list of the data fields or the report form, if possible. A link to a website with the description of the data is also acceptable. If you would like to e-mail the details of the report, please send the file(s) to: kysitlus@tietoanalytics.ee .

25. Which data objects and subjects does the data requested concern? For example, a citizen, vehicle, field, building, etc. *

Please specify about whom or what the data are requested. For example, in the case of the VAT return of a business, data must be provided about the business and its turnover. If the description is available online, please provide the link here.

26. Which classifications or lists are used in the report? *

It is important to specify whether the classifications used by the agency were established at the national or international level (e.g. managed by Statistics Estonia) and which are the classifications compiled and managed by the agency itself. Is there a classification management system which would enable exporting or publishing the classification information for this questionnaire?

27. What is the legal basis for requesting the report? *

Please list the legislative acts or other legal basic documents which legalise requesting the data from the reporting party.

28. Describe a template, form, or data model of the report. *

Add links to the documents/online sources where the data content of the reports is described. If you would like to e-mail a template, form, or data model of the report, please send the file(s) to: kysitlus@tietoanalytics.ee.

29. Describe the needs of the agency to develop the report. *

List the shortcomings of submitting the reports and processing the data and describe any further need for reporting. Should this reporting be terminated or handed over to another agency? How could the reporting-related work of the agency be simplified with respect to this report? What should the state do to develop reporting?

30. Which reporting-related developments is the agency planning in the near future? *

31. Which channel do you prefer for sending the report in the future (online questionnaire, X-tee, e-mail, documents enclosed to e-mails, etc.)? *

Please list your preferred solutions for organising reporting.

Your response has been submitted. Thank you! Please click here to submit the next report:

<https://forms.office.com/r/rMSR7GqDLf>

Please click here for an overview of the responses submitted: https://tietoanalytics.ee/MKM_aruanded

Another form for submitting reports 2+

Reporting questionnaire – adding a further report

The purpose of this questionnaire is to support the simplification and standardisation of regular reporting. The questionnaire is designed for the state agencies which receive periodic reports from private companies.

The surveying is organised based on public procurement no. 231015 organised by the Ministry of Economic Affairs and Communications. The contracting authority has authorised Tieto Estonia AS to serve as the processor of the data collected within the framework of the contract.

The questionnaire is concerned with the periodic reporting collected by your agency. Please provide information about all periodic reports.

Responding to the questionnaire takes at least 30 minutes, on average, depending on the level of complexity of the reporting of the agency.

Please address any questions about the questionnaire or responding to the questionnaire to kysitlus@tietoanalytics.ee or call +372 5903 9494.

Information about the reports used at the agency.

Below, you can submit the information about one report used at the agency. Please only provide information about the reports which are submitted periodically (not single reports submitted on one occasion). The descriptions of the reports must be submitted separately for each report. Having filled in the questionnaire about one report, you can start submitting the information of another report without having to fill in the general information block again.

1. What agency are you responding for? *
2. What is the name of the report? *
3. How often are the companies required to submit the report? *
 - a. regularly
 - b. regularly updated
 - c. once per hour
 - d. once in two hours

- e. once in three hours
- f. twice a day
- g. once a day
- h. once a week
- i. three times a week
- j. twice a week
- k. three times a month
- l. twice a month
- m. once a month
- n. once in two months
- o. once a quarter
- p. three times a year
- q. once in six months
- r. once a year
- s. once in two years
- t. once in three years
- u. once in four years
- v. once in five years
- w. once in ten years
- x. once in twenty years
- y. once in thirty years
- z. irregularly
- aa. unknown
- bb. never

cc. Other

4. Which period the report is submitted for? *

a. regularly

b. 1 hour

c. 3 hours

d. 12 hours

e. 1 day

f. 2 days

g. 3 days

h. 1 week

i. 10 days

j. 2 week

k. 15 days

l. 1 month

m. 2 months

n. 1 quarter

o. 4 months

p. 6 months

q. 1 year

r. 2 years

s. 3 years

t. 4 years

u. 5 years

v. 10 years

- w. 20 years
- x. 30 years
- y. irregularly
- z. unknown
- aa. never
- bb. Other...

5. What is the term for submitting the report? *

E.g. Date, month

6. Please provide a brief description of the content of the report. *

Please describe briefly the subject and purpose of submitting the report

7. In which data medium/format can the report be submitted? *

Please specify all options available.

- a. csv
- b. doc
- c. paper
- d. pdf
- e. xbrl-gl
- f. xls(x)
- g. xml
- h. Other

8. Which channel and/or data exchange platform can be used to submit the report? *

Please specify all options available.

- a. X-tee
- b. Other machine interface

- c. Uploading a file
- d. Self-service environment
- e. Email
- f. Regular mail
- g. On-site
- h. Other

9. What is the total number of data fields which must be submitted by the reporting parties? *

The total number of data fields means the number of questions to which the reporting party must respond. If a question includes several sections, it may be counted as one data field (e.g. the amount of fuel consumed in tons, by types of fuel).

10. Which data (data fields) are requested by the report? *

Here, you may provide a list of the data fields or the report form, if possible. A link to a website with the description of the data is also acceptable. If you would like to e-mail the details of the report, please send the file(s) to: kysitlus@tietoanalytics.ee .

11. Which data objects and subjects does the data requested concern? For example, a citizen, vehicle, field, building, etc. *

Please specify about whom or what the data are requested. For example, in the case of the VAT return of a business, data must be provided about the business and its turnover. If the description is available online, please provide the link here.

12. Which classifications or lists are used in the report? *

It is important to specify whether the classifications used by the agency were established at the national or international level (e.g. managed by Statistics Estonia) and which are the classifications compiled and managed by the agency itself. Is there a classification management system which would enable exporting or publishing the classification information for this questionnaire?

13. What is the legal basis for requesting the report? *

Please list the legislative acts or other legal basic documents which legalise requesting the data from the reporting party.

14. Describe a template, form, or data model of the report. *

Add links to the documents/online sources where the data content of the reports is described. If you would like to e-mail a template, form, or data model of the report, please send the file(s) to: kysitlus@tietoanalytics.ee .

15. Describe the needs of the agency to develop the report. *

List the shortcomings of submitting the reports and processing the data and describe any further need for reporting. Should this reporting be terminated or handed over to another agency? How could the reporting-related work of the agency be simplified with respect to this report? What should the state do to develop reporting?

16. Which reporting-related developments is the agency planning in the near future? *

17. Which channel do you prefer for sending the report in the future (online questionnaire, X-tee, e-mail, documents enclosed to e-mails, etc.)? *

Please list your preferred solutions for organising reporting.