

Preliminary study

DATA RE-USE

15 May 2024

Nr. ITE-1

SUMMARY

Relevance of the audit subject matter

Public sector data is a valuable resource whose re-use can generate political, economic and social benefits through digital innovation, new products and services, research, and important data-driven decisions. Lithuania aims to open up this data centrally for re-use by business, academia and public authorities. According to estimates¹ by the Ministry of the Economy and Innovation, data availability and rational re-use would increase the efficiency of the public sector and save the state budget between EUR 0.45 million and EUR 2.34 million in 2025, and between EUR 0.80 million and EUR 4.21 million in 2026. The size of the open data market in Lithuania could reach EUR 125 million in 2025 and EUR 285 million in 2026 and contribute to the creation of new jobs. According to estimates published by the Information Society Development Committee, the use of open data could generate 2% of GDP.

To maximise impact, it is important to facilitate the effective re-use of data by ensuring that open data are of high quality, easily accessible and user-friendly for re-use by the public, the business and scientific sectors, and by private individuals.

¹ Internet access: <https://eimin.lv.lt/uploads/eimin/documents/files/Open%20data%20Poveikis%20ekonomikai.pdf> (accessed on 7 February 2024).

Objective and scope of preliminary study

The aim of the preliminary study² is to collect and assess information on data opening and re-use in Lithuania and to identify problems in the field.

The main questions of the preliminary study were aimed at analysing:

- the data re-use management system;
- the legal framework;
- progress in implementing the strategic goals and objectives;
- the quality of open data;
- the need to re-use data for commercial and non-commercial purposes.

Audited entities:

- The Ministry of the Economy and Innovation, as it formulates state policy on data re-use and organises, coordinates and monitors its implementation;³;
- The Information Society Development Committee, as it coordinates the opening up of data of institutions and state-owned entities, provides them with methodological support and monitors their opening up of data⁴.

We gathered information and communicated with representatives of the Ministry of the Economy and Innovation, the Information Society Development Committee, the State Data Agency, the Government Strategic Analysis Centre, the Central Project Management Agency, the State Security Department of the Republic of Lithuania, the Second Operational Services Department under the Ministry of National Defence and the Lithuanian Business Confederation.

The period covered is 2019-2023. In order to assess trends, we have used data from previous years and, in some cases, developments in 2024.

The preliminary study was carried out in accordance with international standards of supreme audit institutions. The methods used in the preliminary study are described in more detail in Annex 2 "Methods used for data collection and evaluation" (p. 23).

² Preliminary study is a phase in the performance audit process that gathers and evaluates information about the area of activity under examination in order to identify the problems it contains. The results of the preliminary study form the basis for the main study. However, in certain circumstances, in order to rationalise the use of audit resources, the main study may be omitted, and the audit concluded with a preliminary study: e.g. the audited entity has taken action to resolve the problems, there have been changes in the audited area that may prevent further audit work from having the required impact, etc. The preliminary study report identifies the problems found without making recommendations to the audited entity.

³ Law on the right to information and re-use of data, Article 7(2).

⁴ Ibid, Article 7(3), Regulations of the Information Society Development Committee approved by Order No 4-593 of the Minister of the Economy and Innovation of 1 October 2018, p. 12.

Key results of the preliminary study

Užtikrinus rezultatyvų atvirų duomenų iniciatyvų įgyvendinimą, aukštesnį atvertų duomenų rinkinių brandos lygį, veiksmingą duomenų naudotojų pateiktų poreikių vertinimą didės atveriamų duomenų pakartotinio naudojimo galimybės. Sustiprinta atveriamų duomenų peržiūra didins jautrių duomenų saugą.

There is a need to ensure effective implementation of open data initiatives

Since the end of 2016, the Information Society Development Committee has been implementing a project to develop the Lithuanian Open Data Portal, which has been funded with more than EUR 2.577 million. The planned activities were implemented on the 31 October 2023, but as of 12 April 2024, the project has not been completed and the cost of the project has increased by 24% (to EUR 3.197 million) over the 7 years of implementation. Of the financial projections (EUR 8.37 million) foreseen for the 2014-2020 Open Data projects, more than half (almost 62% or EUR 5.17 million) remain unused.

Due to the lack of effectiveness in the implementation of the 2014-2020 open data initiatives, 75% (3 out of 4) of the evaluation criteria set out in the Information Society Development 2014-2020 Programme "Digital Agenda of the Republic of Lithuania" were not achieved in 2020.

When analysing the status of implementation of the activities of the progress measure "Promote data availability and re-use" of the State Digitalisation Development Programme for the period 2021-2030 of the Ministry of the Economy and Innovation, we found that in 2022-2023, 1 out of the 3 projects approved in the new development period are delayed in implementation and 25% (1 out of 4) of the targets were not met (Section 1, page 11).

There are cases where opening up data does not assess the impact on national security

When opening up datasets, public sector entities are responsible for the correctness of the data provided, but, for example, the use of artificial intelligence or machine learning to process datasets and merge bundles of data (including personalised data) can help to design important decisions that can also have a negative impact on Lithuania's national security. Public sector actors opening up data may not be aware that opening up existing data and analysing it using AI algorithms for intelligence purposes may affect Lithuania's national security. The Department of State Security and the Second Department of Operational Services, in their assessment of risks, hazards and threats to Lithuania's national security, identified 2 cases of data openings in 2023 that posed a threat to national

security, but they are not included in the review process of the opened data, leaving the responsibility to the public sector entities (Section 2, p. 15).

Open datasets of insufficient maturity reduce their usability

The possibilities for reuse of open data are limited due to the lower than recommended maturity level. In 2023, 253 public sector entities have published 1 436 datasets on the Lithuanian Open Data Portal, but information on the maturity level of half of them (51%) is not published, 40% are available at maturity levels 3 and 4, 8% are at a lower maturity level, and one in five (20%) are published in an un-open format, limiting the possibilities of re-use of these datasets.

According to data from the Ministry of the Economy and Innovation and the Information Society Development Committee, in 2022, more than 50% of institutions had not inventoried and opened any datasets, and only a third (29%) were planning to open them in the next year.

Looking at open datasets published on other portals, we found that not all of the data published there is available in the one-stop shop i.e. the Lithuanian Open Data Portal.

Data users' needs are not sufficiently assessed. A total of 241 requests for datasets were received over the 4-year period, and a small proportion (22%) were fulfilled. There is no data on the extent to which requests are made by individuals, legal entities, representatives of public and other sectors, and the Information Society Development Committee does not collect or hold this information (Section 3, p. 17).

Reasons for concluding an audit with a preliminary study

In 2022, the Ministry of the Economy and Innovation approved a list of activities⁵ to promote data availability and re-use and started the procedures for their implementation during the first three years of the 2021-2030 development period. The plans are to develop a model for the creation and management of metadata and software interfaces, to ensure the availability of metadata management software, to implement integration interfaces between at least 376 information systems and registers, to create a repository of software interfaces populated with metadata from at least 376 information systems and registers, to integrate the State's information resources into a data lake, to create a data exchange tool, to encourage business investment in the development of innovative new solutions and digital services using open data⁶.

In 2023, the State Data Agency launched the project "Integration of State Information Resources into the State Data Lake". The project prioritises the opening up of high-value

⁵ The description of the progress measure No 05-002-01-07-07 "Promote data availability and re-use", Section III of the State Digitisation Development Programme of the Ministry of Economy and Innovation for the period 2021-2030, approved by Order No 4-864 of the Minister of the Economy and Innovation of 12 July 2022.

⁶ Description of Progress Measure 05-002-01-07-07, Section III.

datasets to be made available for re-use across the EU in accordance with an international standard⁷ as from 9 June 2024. The project aims to open up 376 information systems and registers on the Lithuanian Open Data Portal by 1 June 2026, at a minimum maturity level of 4⁸.

Taking into account the developments in 2023 and the activities planned and implemented by the Ministry of the Economy and Innovation, we see that the problems identified during the audit are being addressed and the audit is therefore concluded with a preliminary study. The insights provided in the report will be useful for the implementation of the planned measures to promote data accessibility and re-use, for decision-making on improving the open data process and for strengthening the review of open data.

⁷ From 9 June 2024, geospatial, meteorological, earth observation and environmental, company and corporate governance ownership, statistical and mobility datasets must be made available in computer-readable formats using APIs and, where appropriate, the possibility of bulk downloading of the datasets is ensured, Art.5, Art.6, Art.8, Art.14(1) of the Implementing Regulation (EU) No 2023/138 of the European Commission adopted on 21 December 2022, which draws up the list of specific high value data sets and lays down procedures for making them available and for their reuse.

⁸ Internet access: <https://vda.lrv.lt/lt/veiklos-sritys/projektai/vykdomi/valstybes-informaciniu-istekliu-integravimas-i-duomenu-ezera/> (accessed on 12 March 2024).