



Estonian positions towards Open data – availability of public datasets

I. General remarks

Estonia supports the initiative for an implementing act to enhance the availability of public datasets. The EU should strive to improve the availability of high-quality public-sector data by expanding the categories of re-usable data, for instance, to language corpora. In order to progress in this regard, the European Commission should develop financial support measures for the Member States. It is important to enable cross-border access to data. The exchange of EU public sector data should be based on common standards, for instance, Data Catalogue vocabulary application profile, which is used for describing public sector datasets. DCAT-AP enables and improves cross-data portal search. Similarly important are EU core vocabularies, which standardise the metadata.

It is important to find ways to achieve even higher availability of public sector data sets made available by the public authorities of the Member States. For example, improving the availability of data requires costly IT developments, while ensuring data quality and sufficient amounts of metadata requires skilled specialists currently lacking in the workforce across Europe. Thus, both financial measures and retraining is required. When changing the legal framework for open data, it must be emphasized that data sharing does not restrict the competitiveness of companies, including state-owned companies. It is also important that public sector data is actually available to all interested parties and such parties have actual access to the said data. This would allow the private sector to build new services using public sector data.

Although one of the aims of the EU Data Strategy is to address the shortcomings in the exchange of data between the public sector institutions and to achieve the implementation of the one-stop-shop principle for public services, there is a need for clarifications. Applying one-stop-shop principle is hampered by the lack of semantic interoperability, differences in data formats, insufficient metadata (data is only described at dataset level, while for the use of data, element level description is crucial) and the inability to bring people and data together. With regard to the "one-stop-shop" principle, there is also a need for an increased dialogue between the Member States and the Commission to ensure a common understanding of its content. The lack of cross-border interoperability of public sector data is an obstacle to the creation of a single EU digital gateway. The cross-border use of data also requires a significant development of language technologies so that large text corpora (eg in the National Library, the National Archives and the National Broadcasting) can be used in the development of services.

The other issue with publicly available datasets and the re-use of such datasets is related to the problems of processing personal data. We provide our initial assessment and concerns in relation to this implementing act and its potential effects on processing of personal data in the following paragraphs.

II. High Value Datasets – issues with personal data and compliance with Regulation (EU) 2016/679

Regarding the list of High Value Datasets (hereinafter **HVD**) related to the company and company ownership category, we would like to point out that proposed list contains numerous personal data capable of identifying natural persons. This brings about the concern how making the data in the list re-usable is compatible with the Regulation (EU) 2016/679 on data protection (hereinafter **GDPR**).

Our concerns are mainly related to following issues which have partially been addressed in the recitals of the Directive (EU) 2019/1024 on open data and the re-use of public sector information. We would like to get the perspective of the Commission on these issues as well as an opinion from the European Data Protection Supervisor to make sure that any future steps taken do not damage fundamental rights of the individuals and go against GDPR principles.

2.1. Permissible re-use of personal data in case of re-use of publicly available data

In recital 52 of the Directive (EU) 2019/1024 on open data and the re-use of public sector information (hereinafter **Open Data Directive**), among others, following has been stated:

“This Directive does not affect the protection of individuals with regard to the processing of personal data under Union and national law, particularly Regulation (EU) 2016/679 and Directive 2002/58/EC of the European Parliament and of the Council (18) and including any supplementing provisions of national law. This means, inter alia, that the re-use of personal data is permissible only if the principle of purpose limitation as set out in point (b) of Article 5(1) and Article 6 of Regulation (EU) 2016/679 is met.”

This recital emphasises the importance of respecting purpose limitation set out in point (b) of Article 5(1) of Regulation (EU) 2016/679. For example, the legal basis for Estonian business register is established in the Commercial Code (hereinafter **CC**). § 22 of CC stipulates that purpose of maintaining such database is to collect, store and disclose information on the enterprises of sole proprietorships, companies and branches of foreign companies located in Estonia. The data, including personal data, is held in business register until such purpose exists after which the data will be deleted under § 59 of CC.

However, if this data has been shared in machine readable format and has been multiplied by different service providers, the data subject cannot be certain that this not up to date data does not still exist in other privately held databases. This would violate not only the purpose limitation principle, but would breach the principles of lawful processing and prohibition to keep personal data in a form which permits identification of data subjects for longer than necessary for the purposes for which the personal data are processed GDPR Article 5(1) points (a) and (e). This also applies to the changes made in the business register and how to ensure that the information about the relation of a natural persons to a specific company is up to date in different private databases or services based on business register data. Of course, data subject always has the right to turn to a private body processing his/her data and to ask data to be deleted or amended, however it would bring about considerable burden for natural persons to handle separately all the privately held

databases or services. This possibly weakens the level of protection provided to natural persons vis á vis processing their personal data.

Furthermore, if we include accounting documents, such as company annual reports, then in the case of micro enterprises (one-man companies which Estonia has more than 120,000) information about the income of a natural person and possibly economic situation can be deduced from the data in such reports. We understand (from GDPR recital 71) that this is normally seen as data that needs to be processed with more care and precautionary measures than some other categories of personal data.

2.2. Profiling of natural persons based on public sector data of public registries

We would also like to enquire if the Commission and European Data Protection Supervisor see a possibility that personal data contained in public registries (for example the business registry) could be used, in combination with other publicly available data, as part of profiling natural persons. There are already vast amounts of personal data available online and these amounts are expected to only grow in the future. This could possibly hinder natural persons from enforcing their fundamental rights effectively, especially in the situations where private databases containing personal data from different public sources have been created. The solution that will be foreseen needs to be future proof vis á vis protecting personal data and fundamental rights.

In case of the business registry, for example in relation to one-person companies, if some sets of data are processed together, such as company ownership information in the case of micro enterprises combined with the annual reports, can give an adequate overview of a person and his financial situation, which in turn can, in some specific cases, lead to decisions where profiling has been used.

In conclusion, enabling a wider use and re-use of public datasets is encouraged, however, first the aforementioned issues in relation to GDPR need to be addressed to avoid a situation where the implementing act does not comply with the rules of an EU Regulation.