

PUBLICATION OF HIGH-VALUE DATASETS UNDER THE ACT NO. 106/1999 COLL., ON FREE ACCESS TO INFORMATION ("THE INFORMATION ACT")

ŠKUREK, Martin¹⁴⁵

Introduction

At the end of last year it was reported in the media that the European Commission had released a list of high-value datasets. High-value datasets can be defined as documents the re-use of which is associated with important benefits for society, the environment and the economy.¹⁴⁶ The issue of high-value datasets is thus very topical and important, and the author of the present article is therefore interested in finding out how this legal institute is enshrined in the legal system of the Czech Republic, as one of the EU Member States. The instrument of this enshrinement shall be Amendment no. 241/2022 Coll. which, with effect from January 1st 2023, supplemented and changed certain provisions of the Act no. 106/1999 Coll., on free access to information (hereinafter referred to as "the Information Act").

With regard to its content, the present article deals with the part of the Amendment no. 241/2022 Coll. which, with effect from 1 January 2023, supplemented the legal regulation of the Information Act with a provision (the new Section 5b) on the publication of high-value datasets. The new Section 5b of the Information Act states that the so-called obliged entity must publish as open data, in a machine-readable and open format under standard conditions of use in particular through an application programming interface, the information (i.e. high-value datasets) listed in a directly applicable regulation of the European Union issued under Article 14(1) of the Directive (EU) 2019/1024 of the European Parliament and of the Council on open data and the re-use of public sector information (hereinafter referred to as "the Directive (EU) no. 2019/1024").¹⁴⁷

¹⁴⁵ Faculty of Law, Palacký University in Olomouc Department of Administrative Law and Financial Law

¹⁴⁶ Point 66 of the preamble to the Directive (EU) no. 2019/1024.

¹⁴⁷ On some of the problems associated with the adoption of the Directive (EU) no. 2019/1024, see Vuckovic, R., Kanceljak, I. & Juric, M. (2021). Is It Time for New Data Governance? *2021 44th International Convention on Information, Communication and Electronic Technology, MIPRO 2021 – Proceedings*, pp. 1483–1489. DOI: 10.23919/MIPRO52101.2021.9597161.



Terms linked to the concept of high-value datasets

It follows from Section 5b of the Information Act that the concept of a high-value datasets is linked to other terms, namely the terms of obliged entity, open data, machine-readable format, open format and application programming interface. Obligated entities which are obliged to provide information related to their competence according to Section 2 of the Information Act are state authorities, local self-government units and their bodies and public institutions. Obligated entities are also those entities entrusted by law with decision-making power on the rights, legally protected interests or obligations of natural or legal persons in the field of public administration, only to the extent of their decision-making activities. According to the Information Act, professional self-governing chambers publish only information relating to the exercise of public administration entrusted to them by law. Furthermore, the obliged entity is a public company which for the purposes of Section 2a of the Information Act means a legal entity which (a) carries out a relevant activity under the Act no. 134/2016 Coll., Public Procurement Act; acts as a public service provider under Article 2 of the Regulation (EC) no. 1370/2007 of the European Parliament and of the Council on public passenger transport services by rail and by road and on repealing the Council Regulations (EEC) nos. 1191/69 and 1107/70; acts as an air carrier which performs public service obligations under Article 16 of the Regulation (EC) no. 1008/2008 of the European Parliament and of the Council on common rules for the operation of air services in the Community; or acts as the owner of a ship which discharges a public service obligation under Article 4 of the Council Regulation (EEC) no. 3577/92 applying the principle of freedom to provide services to maritime transport within Member States (maritime cabotage); and (b) in which the obliged entity may exercise, directly or indirectly, a dominant influence by virtue of its ownership interest in the legal person or the rules governing it. The condition of dominant influence shall be deemed to be fulfilled if the obliged entity directly or indirectly (a) holds a majority of the subscribed share capital of the public company, (b) holds a majority of the voting rights attaching to the shareholding in the public company; or (c) may appoint more than half of the members of the administrative, management or supervisory body of the public company.¹⁴⁸

For the purposes of the Information Act, open data means information published in a manner allowing remote access in an open and machine-readable format, the manner and purpose of subsequent use of which is not restricted by the obliged entity publishing it and which is registered in the National Catalogue of Open Data (hereinafter referred to as the "NCOD"). In other words, open data is a highly effective way of publishing public sector information. In particular, open data must be: accessible as data files in a machine-

¹⁴⁸ Section 2 and 2a of the Act no. 106/1999 Coll., on free access to information.

readable and open format with complete and up-to-date database content or aggregated statistics; accompanied by non-restrictive terms of use; registered in the NCOD by obliged entities as direct links to data files; accompanied by documentation; available for download without technical barriers (registration, access restrictions, CAPTCHA, etc.); prepared with the aim of the easiest possible machine processing by programmers, etc. and with contact details for the curator for feedback (bugs, extension requests, etc.). If a dataset does not meet one of the conditions, it is not open data. On the other hand, the following are not open data: a search form for end users, a web page with additional information, a public GIS map interface, a web service making available the individual records that make up a dataset with a limitation on the number of accesses, a dataset in PDF format, a dataset in XLS format for printing or a dataset in pseudo-CSV format (e.g. a separator other than a comma). Open data does not only refer to traceability data but to all areas of public administration.¹⁴⁹

As regards the registration of open data by obliged entities in the NCOD, it is a public administration information system accessible by remote access and used to register information published as open data and other information published by remote access, in particular by an interface for programming applications and by information on products and services that use open data and other information. The information contained in the NCOD shall be provided in a way that allows remote access as open data. The NCOD is managed by the Digital and Information Agency (hereinafter referred to as the "the DIA"). The DIA is the Czech central administrative authority for electronic identification, trust services and public administration information systems. It was established on January 1st 2023 and took over full competence from April 1st 2023. It is to take over the administration of the basic registers, the administration of Czech POINTs (Czech POINT is a Czech state project under which municipal offices with extended competence, regional offices, notaries and other legal entities, e.g. establishments of the Czech Post and local offices of the Chamber of Commerce of the Czech Republic with the appropriate authorization, can issue extracts from the Land Register, from the Criminal Register or from the Trade Register) and the Citizen's Web Portal from the Ministry of the Interior. The DIA was established by the Amendment no. 471/2022 Coll. to the Act no. 12/2020 Coll., The Digital Service Right Act.¹⁵⁰

For the purposes of the Information Act, a machine-readable format means a format of a data file with a structure that allows software to easily find, recognise and extract specific information from that data file, including individual data and their internal structure. In other words, a machine-readable format is one that

¹⁴⁹ Section 3a of the Act no. 106/1999 Coll., on free access to information and The Government of the Czech Republic. Co jsou otevřená data? Open data portal. Available at: <https://opendata.gov.cz/informace:start>. (cited: 30/08/2023)

¹⁵⁰ Section 4c of the Act no. 106/1999 Coll., on free access to information and details of the DIA in the explanatory memorandum of Amendment no. 471/2022 Coll.

can be read by a computer program. Usually, if a public administration worker scans text in image format (JPEG), the information seeker has no problem reading the text, but a machine cannot. The obliged entity must therefore choose between machine-readable formats. The most suitable machine-readable formats are considered to be RDF, Odata, XML, CSV. On the other hand, PDF, DOC or DOCX are not suitable formats.¹⁵¹

Furthermore, for the purposes of the Information Act, an open format means a format of a data file that is not dependent on specific technical and software equipment and is made available to the public without any restriction that would prevent the use of the information contained in the data file. In other words, an open format is one that is accessible (readable) by the requester of the information using common computer software without legal restrictions on use. Thus, an open format can be understood as a format without legal restrictions on use which aims to guarantee long-term access to data without current or future uncertainty about legal rights or technical specifications. In practice, obliged entities provide information in open formats such as TXT, PDF, RTF, RDF, XML, HTML/XHTML, JPEG or CSV. Thus, as can be seen from the above, not every open format is machine-readable, see e.g. PDF format. Obligated entities must therefore consider carefully which format they choose in a given case.¹⁵²

Finally, for the purposes of the Information Act, an application programming interface (hereinafter referred to as "API") means a set of functions and processes that enable the creation of applications that access the functions or data of an operating system, database, application or other electronic service. In other words, API is an interface that is used to communicate between two programs/applications. Through it, they can exchange open data and extend their functionality more easily, for example, with open data that is prepared by someone else. Open data can be freely processed and presented in a web or mobile application via API. API is sometimes thought of as contracts with terms between two parties: if one party sends a remote request with a certain structure, the other party's software will respond. Within the framework of APIs of digital services of the state, APIs of data boxes or APIs of the value-added tax register have been established in the Czech Republic.¹⁵³

¹⁵¹ Section 3a of the Act no. 106/1999 Coll., on free access to information and BUDÍL, P.: „Novela „Infizákona“ – malá revoluce ve zpracování informací?“ *Právní prostor*, 2015, available at: <https://www.pravniprostor.cz/clanky/spravni-pravo/novela-infozakona-mala-revoluce-ve-zpracovani-informaci>. (cited: 30/08/2023)

¹⁵² Ibidem.

¹⁵³ Section 3a of the Act No. 106/1999 Coll., on free access to information and the Ministry of the Interior of the Czech Republic. *Popis API*. Open data of the Ministry of the Interior of the Czech Republic. Available at: <https://opendata.mvcr.cz/popis-api>. (cited: 30/08/2023)

The concept of high-value data

As already indicated, unlike the terms defined above, the Information Act does not explicitly define the concept of high-value datasets, but only refers to the Directive (EU) no. 2019/1024. Point 66 of the preamble to the Directive (EU) no. 2019/1024 states that in order to provide for conditions supporting the re-use of documents which is associated with important socioeconomic benefits having a particular high value for the economy and society, a list of thematic categories of high-value datasets should be set out in an Annex to this Directive. By way of illustration, and without prejudice to the implementing acts identifying the high-value datasets to which the specific requirements set out in this Directive should apply, taking into account the Commission guidelines on recommended standard licences, datasets and charging for the reuse of documents, the thematic categories could, inter alia, cover postcodes, national and local maps (geospatial), energy consumption and satellite images (earth observation and environment), in situ (the original position) data from instruments and weather forecasts (meteorological), demographic and economic indicators (statistics), business registers and registration identifiers (companies and company ownership), road signs and inland waterways (mobility).¹⁵⁴

Under Article 13(1) of the Directive (EU) no. 2019/1024, in order to provide for conditions to support the re-use of high-value datasets, a list of thematic categories of such datasets is set out in Annex I to this Directive. Under Annex I the list of thematic categories of high-value datasets is as follows: 1. Geospatial, 2. Earth observation and environment, 3. Meteorological, 4. Statistics, 5. Companies and company ownership, 6. Mobility. Article 13(2) then stipulates that the Commission is empowered to adopt delegated acts in accordance with Article 15 in order to amend Annex I by adding new thematic categories of high-value datasets in order to reflect technological and market developments.¹⁵⁵

Article 14(1) stipulates that the Commission shall adopt implementing acts laying down a list of specific high-value datasets belonging to the categories set out in Annex I and held by public sector bodies and public undertakings among the documents to which this Directive applies. Such specific high-value datasets shall be: (a) available free of charge, subject to Article 14(3), 14(4) and 14(5); (b) machine readable; (c) provided via APIs; and (d) provided as a bulk download, where relevant. Those implementing acts may specify the arrangements for the publication and re-use of high-value datasets. Such arrangements shall be compatible with open standard licences. The arrangements may include terms applicable to re-use, formats of data and metadata and technical arrangements for dissemination. Investments made by the Member States in open

¹⁵⁴ Point 66 of the preamble to the Directive (EU) no. 2019/1024.

¹⁵⁵ Article 13 of the Directive (EU) no. 2019/1024.

data approaches, such as investments into the development and roll-out of certain standards, shall be taken into account and balanced against the potential benefits from inclusion in the list. Those implementing acts shall be adopted in accordance with the examination procedure referred to in Article 16(2).¹⁵⁶

Under Article 14(2), the identification of specific high-value datasets pursuant to Article 14(1) shall be based on the assessment of their potential to: (a) generate significant socioeconomic or environmental benefits and innovative services; (b) benefit a high number of users, in particular SMEs; (c) assist in generating revenues; and (d) be combined with other datasets. For the purpose of identifying such specific high-value datasets, the Commission shall carry out appropriate consultations, including at expert level, conduct an impact assessment and ensure complementarity with existing legal acts, such as Directive (EU) no. 2010/40/EU on the framework for the deployment of Intelligent Transport Systems in the field of road transport and for interfaces with other modes of transport, with respect to the re-use of documents. With regard to high-value datasets held by public undertakings, the impact assessment shall give special consideration to the role of public undertakings in a competitive economic environment.¹⁵⁷

Article 14(3) states that by way of derogation from point (a) of the second subparagraph of Article 14(1), the implementing acts referred to in Article 14(1) shall provide that the availability of high-value datasets free of charge is not to apply to specific high-value datasets held by public undertakings where that would lead to a distortion of competition in the relevant markets. Under Article 14(4), the requirement to make high-value datasets available free of charge pursuant to point (a) of the second subparagraph of Article 14(1) shall not apply to libraries, including university libraries, museums and archives. And Article 14(5) finally states that where making high-value datasets available free of charge by public sector bodies that are required to generate revenue to cover a substantial part of their costs relating to the performance of their public tasks would lead to a substantial impact on the budget of the bodies involved, Member States may exempt those bodies from the requirement to make those high-value datasets available free of charge for a period of no more than two years following the entry into force of the relevant implementing act adopted in accordance with Article 14(1).¹⁵⁸

Following on from the Directive (EU) no. 2019/1024, the Commission had adopted an implementing act focused on high-value datasets on December 21st 2022 (hereinafter referred to as "the Implementing Regulation"), thus the Implementing Regulation laying down a list of specific high-value datasets and the arrangements for their publication and re-use. Since defining the specific value of datasets was not a trivial

¹⁵⁶ Article 14 of the Directive (EU) no. 2019/1024.

¹⁵⁷ Article 14 of the Directive (EU) no. 2019/1024.

¹⁵⁸ Article 14 of the Directive (EU) no. 2019/1024.

task, an impact assessment study (Impact Assessment study on the list of High Value Datasets to be made available by the Member States under the Open Data Directive) was prepared for the Commission, detailing the list of high-value datasets that were to be made available.¹⁵⁹ The starting point of the study was a map of all relevant EU legislation, presenting the datasets that were already available from all EU Member States. Interviews with stakeholders then took place to develop a preliminary wish list of datasets considered to be of the highest value from an economic and social reuse perspective. Moreover, the Commission provided the inception impact assessment (Open data – availability of public datasets), aiming to inform citizens and relevant stakeholders about its plans and open up to feedback. The document stressed the importance of high-value datasets and the need to have harmonisation rules to improve the availability of public data and its reuse.¹⁶⁰ As a result of this process, a limited and well-defined group of datasets were identified. These aim to provide maximum value to their users and shall be able to be used free of any technical, legal or financial barrier. These datasets are listed in the Implementing Regulation and are in accordance with the Directive (EU) no. 2019/1024 grouped in a list of six high-value datasets thematic categories: geospatial, earth observation and environment, meteorological, statistics, companies and company ownership and mobility.¹⁶¹

In the above context, it should be added that the first guidelines concerning public sector information in the EU were produced in 1989, and since then several policy documents, studies and further legislation had followed. More specifically, public sector information was regulated by the first directive in 2003 (the Directive no. 2003/98/EC of the European Parliament and of the Council on the re-use of public sector information), by the directive in 2007 (the Directive no. 2007/2/EC of the European Parliament and of the Council establishing an Infrastructure for Spatial Information in the European Community, so called INSPIRE),¹⁶² by the directive in 2013 (the Directive no. 2013/37/EU of the European Parliament and of the Council amending the Directive no. 2003/98/EC on the re-use of public sector information), by the general data protection regulation in 2016 (the Regulation (EU) no. 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing the Directive no. 95/46/EC /General Data Protection

¹⁵⁹ The European Commission. *Impact Assessment study on the list of High Value Datasets*. The official portal for European data. Available at: <https://data.europa.eu/en/publications/datastories/high-value-datasets-overview-through-visualisation>. (cited: 30/08/2023)

¹⁶⁰ The European Commission. *Open data – availability of public datasets*. The official portal for European data. Available at: https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12111-Open-data-availability-of-public-datasets_en. (cited: 30/08/2023)

¹⁶¹ The Implementing Regulation to the Directive (EU) no. 2019/1024.

¹⁶² For lessons learned during the implementation of the INSPIRE, see for example KOISEV, A.-PEETERS, O.-SMITS, P.-GROTHE, M.: *Building bridges: experiences and lessons learned from the implementation of INSPIRE and e-reporting of air quality data in Europe* in *Earth Science Informatics*, Vol.2 2015.pp, 353 – 365. DOI: 10.1007/s12145-014-0160-8.

Regulation/) and finally by the latest directive in 2019 (the Directive (EU) no. 2019/1024). The public sector information directives were instrumental in harmonising the public sector information available to the public, increasing transparency and introducing a set of measures (such as the use of machine-readable formats or central repositories) to facilitate the discovery and reuse of information produced by the public administration. This means that the Implementing Regulation establishing high-value datasets is the culmination of a process developed over many years.

Specifically, its Article 1 states that the Implementing Regulation establishes the list of high-value datasets belonging to the thematic categories set out in Annex I to the Directive (EU) no. 2019/1024 and held by public sector bodies among the existing documents to which this Directive applies. The Implementing Regulation also lays down the arrangements for publishing and reusing high-value datasets, in particular the applicable conditions for re-use and the minimum requirements for disseminating data via API.¹⁶³

Article 3 of the Implementing Regulation stipulates that public sector bodies holding high-value datasets listed in its Annex shall ensure that the datasets described or referenced in its Annex are made available in machine-readable formats via APIs corresponding to the reasonable needs of re-users. Where indicated in the Annex to the Implementing Regulation, the datasets shall also be made available as a bulk download. Those public sector bodies referred to in Article 3(1) of the Implementing Regulation shall set out and publish the terms of use of the API and the quality of service criteria on its performance, capacity and availability. The terms of use shall be available in a human-readable and machine-readable format. Both the terms of use and the quality of service criteria shall be compatible with the arrangements for the re-use of high-value datasets laid down in accordance with Article 4 of the Implementing Regulation. API terms of use shall be accompanied by API documentation in the European Union or internationally recognised open, human-readable and machine-readable format. Public sector bodies referred to in Article 3(1) of the Implementing Regulation shall designate a point of contact for questions and issues related to the API with a view to ensure the availability and maintenance of the API and ultimately the smooth and effective publication of the high-value datasets. Public sector bodies holding high-value datasets listed in the Annex to the Implementing Regulation shall ensure that the datasets are denoted as high-value datasets in their metadata description.¹⁶⁴

Article 4 of the Implementing Regulation states that an exemption granted by a Member State in accordance with Article 14(5) of the Directive (EU) no. 2019/1024 shall be published online, in the same way as the list of public sector bodies referred to in Article 6(3) of the Directive (EU) no. 2019/1024. To facilitate the

¹⁶³ Article 1 of the Implementing Regulation to the Directive (EU) no. 2019/1024.

¹⁶⁴ Article 3 of the Implementing Regulation to the Directive (EU) no. 2019/1024.

availability of datasets for re-use covering longer periods of time, the obligations imposed under the Implementing Regulation shall also apply to existing machine-readable high-value datasets created before the date of application of the Implementing Regulation. High-value datasets shall be made available for re-use under the conditions of the Creative Commons Public Domain Dedication (CC0) or, alternatively, the Creative Commons BY 4.0 licence or any equivalent or less restrictive open licence, as set out in the Annex to the Implementing Regulation, allowing for unrestricted re-use. The Creative Commons BY 4.0 licence is a type of the Creative Commons license. This is a non-copy left free license that is good for art and entertainment works, and educational works. It is compatible with all versions of the GNU, i.e. General Public License. A requirement of attribution, giving the credit to the licensor, can additionally be required by the licensor. High-value datasets shall be made available in accordance with the arrangements for the publication and re-use as set out in the Annex to the Implementing Regulation.¹⁶⁵

Article 5 of the Implementing Regulation finally states that by 2 years after entry into force of this Implementing Regulation Member States shall provide the European Commission with a report on the measures they have carried out to implement the Implementing Regulation. Where appropriate, the information in the report can be provided through references to relevant metadata. The report shall contain the following information: a) a list of specific datasets at Member State level (and, where relevant, subnational level) corresponding to the description of each high-value dataset in the Annex to Implementing Regulation and with online reference to metadata that follow existing standards, such as a single register or open data catalogue; b) persistent link to the licensing conditions applicable to the re-use of high-value datasets listed in the Annex to Implementing Regulation, per dataset referred to in point a); c) persistent link to the APIs ensuring access to the high-value datasets listed in the Annex to Implementing Regulation, per dataset referred to in point a); d) where available, guidance documents issued by the Member State on publishing and reusing their high-value datasets; e) where available, the existence of data protection impact assessments carried out in accordance with Article 35 of the Regulation (EU) no. 2016/679; f) the number of public sector bodies exempted in accordance with Article 14(5) of the Directive (EU) no. 2019/1024.¹⁶⁶ According to the Implementing Regulation, the list of high-value datasets is as follows:

1. Geospatial – The geospatial thematic category includes datasets, within the scope of the INSPIRE data themes titled Administrative units, Geographical names, Addresses, Buildings and Cadastral parcels, as defined in Annex I to the Directive no. 2007/2/EC as well as data themes titled Reference parcels and Agricultural parcels as defined in the Regulation (EU) no. 1306/2013 of the European Parliament and of

¹⁶⁵ Article 4 of the Implementing Regulation to the Directive (EU) no. 2019/1024.

¹⁶⁶ Article 5 of the Implementing Regulation to the Directive (EU) no. 2019/1024.

the Council on the financing, management and monitoring of the common agricultural policy and in the Regulation (EU) no. 1307/2013 of the European Parliament and of the Council establishing rules for direct payments to farmers under support schemes within the framework of the common agricultural policy and the related delegated and implementing acts, setting out their granularity and geographical coverage and listing some or all of the key attributes.

2. Earth observation and environment – The earth observation and environmental thematic category includes earth observation, including space-based or remotely-sensed data, as well as ground-based or in situ (the original position) data, environmental and climate datasets, within the scope of the INSPIRE data themes, as defined in Annexes I-III to the Directive no. 2007/2/EC and datasets produced or generated in the context of the relevant legal acts. The most up-to-date datasets as well as historical datasets available at all generalisation levels up to the level of scale 1:5000 covering the entire country when combined are included. Furthermore, the earth observation and environmental thematic category includes all Environmental information, as defined in Article 2 of the Directive no. 2003/4/EC of the European Parliament and of the Council on public access to environmental information and repealing Council Directive 90/313/EEC, and the environmental information listed in Article 7 of the Directive no. 2003/4/EC titled Dissemination of environmental information.

3. Meteorological – The meteorological thematic category includes datasets on observations data measured by weather stations, validated observations (climate data), weather alerts, radar data and numerical weather prediction model data with the granularity and the relevant key attributes.

4. Statistics – The statistics thematic category includes statistical datasets related to reporting obligations as defined in the relevant legal acts, i.e. Industrial production, Industrial producer price index breakdowns by activity, Volume of sales by activity, EU International trade in goods statistics – exports and imports breakdowns simultaneously by partner, product and flow, Tourism flows in Europe, Harmonised Indices of consumer prices, National accounts – GDP main aggregates, National accounts – key indicators on corporations, National accounts – key indicators on households, Government expenditure and revenue, Consolidated government gross debt, Environmental accounts and statistics, Population, Fertility, Mortality, Current healthcare expenditure, Poverty, Inequality, Employment, Unemployment, Potential labour force. In order to identify in a complete manner the relevant legal references in the relevant legal acts, for some of the datasets it is necessary to refer to concepts from a number of provisions and Annexes. Time series shall begin no later than at the date of application of the relevant respective legal act.

5. Companies and company ownership – The companies and company ownership thematic category includes datasets containing basic company information and company documents and accounts at

individual company level and with the relevant key attributes, i.e. Name of the company, Company status, Registration date, Registered office address, Legal form, Registration number, Member State where the company is registered and Activity/activities that are the object of the company, such as the NACE (Nomenclature statistique des Activités économiques dans la Communauté Européenne) code.

6. Mobility – The mobility thematic category includes datasets, within the scope of the INSPIRE directive data theme titled Transport networks, as set out in Annex I to the Directive no. 2007/2/EC, at all generalisation levels available up to a scale of 1:5000, covering the entire country when combined, including as key attributes national identification code, geographical position as well as links with cross-border networks, where available. For those Member States to which the Directive no. 2005/44/EC applies, this category also includes the relevant datasets, i.e. a) Static data – Fairway characteristics, Long-time obstructions in the fairway and reliability, Rates of waterway infrastructure charges, Other physical limitations on waterways, Regular lock and bridge operating times, Location and characteristics of ports and transshipment sites, List of navigation aids and traffic signs, Navigation rules and recommendations; b) Dynamic data – Water depths contours in the navigation channel, Temporary obstructions in the fairway, Present and future water levels at gauges, State of the rivers, canals, locks and bridges, Restrictions caused by flood and ice, Short term changes of lock and bridge operating times, Short term changes of aids to navigation; c) Inland electronic and navigational charts – Waterway axis with kilometres indication, Links to the external xml-files with operation times of restricting structures, Location of ports and transshipment sites, Reference data for water level gauges relevant to navigation, Bank of waterway at mean water level, Shoreline construction, Contours of locks and dams, Boundaries of the fairway/navigation channel, Isolated dangers in the fairway/navigation channel under and above water, Official aids-to-navigation, for example buoys, beacons, lights, notice marks. The datasets are understood as described in the Directive no. 2005/44/EC.¹⁶⁷

Evaluation of the reflection of the legal institute of high-value datasets in the legal system and reality of the Czech Republic

Regarding the reflection of the involvement of the legal institute of high-value datasets in the Czech Republic it is necessary to distinguish between de facto and de jure situation while the European Commission stipulated that EU Member States must make a list of high-value datasets available free-of-charge by June 2024.¹⁶⁸ The de facto situation in each EU Member State is the subject of the Open Data Maturity Report.

¹⁶⁷ Annex of the Implementing Regulation to the Directive (EU) no. 2019/1024.

¹⁶⁸ For information on the situation in each EU Member State, see for example

This report was prepared as part of data.europa.eu project that is an initiative of the European Commission. Regarding the situation in the Czech Republic, the report states that one of the main tasks is to prepare an open data publication list (plan) with selected datasets with identified risks and benefits, along with a time framework for publication. The Czech Republic shall enable individual datasets, including those labelled as high-value datasets, to be directly linked to the list of reuse examples on the national Open Data Portal (which shall also be in open data format). For that reason, when reading the metadata for datasets that shall be labelled as high-value datasets, it shall be possible to see examples of their practical reuse. The national Open Data Portal is also important because metadata quality measurements, downloadable as CSV files, are displayed on it. Moreover, the metrics that are used to measure quality are included in the standard query language and protocol (SPARQL) endpoint of the portal and are part of the user interface for each individual metadata item in the portal. Furthermore, a dashboard to check metadata quality is available for the metadata publishers and users of the portal to better understand potential deficiencies. The Czech Republic also leverages other methods to grasp users' needs, such as surveys and uses the categories related to the type of service provided, like applications, web-services and others. In the Czech Republic several major projects are being implemented. For example the project "Map of educational failure/success" focuses on the links between the level of education and social issues, raising awareness on regional education inequalities that need to be addressed. The Czech Republic shall also run a project called "National Environmental Reporting Platform" focusing on available sources of environmental open data and analysing their impact on the social, political and legislative requirements caused by the climate and environmental changes. In connection with the above, the national open data team provides several means of support to the publication process. On the basis of the above mentioned facts, it can be concluded that in the Czech Republic at the de facto level fundamental processes are being implemented and some effective tools are being used to integrate the legal institute of high-value datasets, enshrined in the Directive (EU) no. 2019/1024, into social life.¹⁶⁹

However, the de jure situation in the Czech Republic is much more complicated. The reason is the fact that the special part of the explanatory memorandum to Amendment no. 241/2022 basically states what is briefly set out in Section 5b of this Amendment.¹⁷⁰ Only the fragment in the special part of the explanatory memorandum of Amendment no. 241/2022 Coll. which states that Member States should publish the

DE JUANA-ESPINOSA, S. - LUJÁN-MORA, S.: *Open government data portals in the European Union: Considerations, development, and expectations* in: *Technological Forecasting and Social Change*, 2019 p.149, DOI: 10.1016/j.techfore.2019.119769;

DE JUANA-ESPINOSA, S. - LUJÁN-MORA, S. (2020). *Open government data portals in the European Union: A dataset from 2015 to 2017 in Data in Brief*, Vol.4, 2020, pp. 105–156. DOI: <https://doi.org/10.1016/j.dib.2020.105156>; or

RIVZA, B.-RIVZA, P.: *Digitalization in the Baltic States*. *International Multidisciplinary Scientific GeoConference Surveying Geology and Mining Ecology Management*, SGEM 2020, pp. 135–144. DOI: 10.5593/sgem2020/2.1/s07.018.

¹⁶⁹ The European Commission: *The Open Data Maturity Report 2022. The official portal for European data*. Available at: <https://data.europa.eu/en/news-events/news/open-data-maturity-report-2022-out>. (cited: 30/08/2023)

¹⁷⁰ The explanatory memorandum of the Amendment no. 241/2022.

information from the datasets in such a way that, with the exceptions provided for in Article 14(3), (4) and (5) of the Directive (EU) no. 2019/1024, is freely available, machine-readable, accessible through an application programming interface and, where relevant, also available with the option to download the complete dataset, while the proposed Section 5b implements these requirements in line with the Information Act system, is essential. A specific part of the explanatory memorandum to the Amendment no. 241/2022 Coll. thus speaks about the implementation of the Directive (EU) no. 2019/1024 in relation to the content of this Amendment. The concept of implementation of the EU directive is often confused with the concept of transposition of the EU directive. However, it should be noted that these are not identical concepts.¹⁷¹ Transposition is merely the formal incorporation of the EU directive into the national legislation (transposition measure) of the Member State, whereas implementation of the EU directive is a broader process, as it involves both transposition and the phase following transposition, i.e. ensuring adequate application of the transposition measure and effective enforcement of the rights and obligations arising from the EU directive.¹⁷²

Thus, if the new Section 5b of the Information Act states that the so-called obliged entity must publish as open data, in a machine-readable and open format under standard conditions of use, in particular through API, the information listed in a directly applicable regulation of the European Union issued under Article 14(1) of the Directive (EU) no. 2019/1024, it is certainly not possible to say that the new Section 5b of the Information Act is, regarding high-value datasets, an effective instrument for the implementation of this Directive. The new Section 5b of the Information Act, regarding high-value datasets, provides no instrument for effective enforcement of the rights arising therefrom. Regarding high-value datasets, the new Section 5b of the Information Act can therefore only be assessed as the result of a completely formal transposition of the Directive (EU) no. 2019/1024, or it is only a reference to Article 14(1) of this Directive. In view of this fact, it is possible to question the meaningfulness of the existence of the new Section 5b of the Information Act, which is constructed in this way. It is also questionable to what extent the new Section 5b of the Information Act fulfils Article 288 of the Treaty on the Functioning of the European Union. Article 288 of the Treaty on the Functioning of the European Union specifically states that a EU directive shall be binding,

¹⁷¹ The Government of the Czech Republic. *Usnesení vlády České republiky ze dne 15. listopadu 1999 č. 1212 + P o Metodických pokynech pro zajišťování prací na sblížování právních předpisů České republiky s právem Evropských společenství*. Informační systém Opatření Úřadu vlády České republiky. Available at: https://kormoran.odok.cz/usneseni/usneseni_webtest.nsf/0/B7724FA12D7EF030C12571B6006D7D54. (cited: 30/08/2023)

¹⁷² For examples of timely and correct transposition of EU directives, see STAROŇOVÁ, K. & KIŠDURKA, M.: *Transposition in new member states of the EU (Transpozícia v nových členských krajinách EÚ)* (2013) in *Sociologia (Slovakia)* Vol. 1 2013, pp. 48 – 70. and on the issue of transposition of the EU directives using Germany as an example, see further PAASCH, J.-STECKER, C.: *When Europe hits the subnational authorities: The transposition of EU directives in Germany between 1990 and 2018* in *Journal of Public Policy*, Vol. 4-2021. pp. 798–817. DOI: 10.1017/S0143814X20000276.

as to the result to be achieved, upon each Member State to which it is addressed, but shall leave to the national authorities the choice of form and methods. As mentioned above, the new Section 5b of the Information Act apparently fails to establish the appropriate forms and methods. Overall, it can be summarized that the new Section 5b of the Information Act does not provide a sufficient platform for the obliged entities and these entities have to follow, basically in its entirety, only the relevant provisions of the Directive (EU) no. 2019/1024. For the aforementioned reasons, the new Section 5b of the Information Act cannot be considered as a functional national regulation for the publication of high-value datasets.

Conclusion

To conclude this article, it is unfortunately necessary to state that even the Directive (EU) no. 2019/1024 may not have a fully effective impact on the publication of high-value datasets in the Czech Republic. It is certainly positive that Article 14(1) and (2) of the Directive (EU) no. 2019/1024 sets out quite clearly the conditions for the publication of high-value datasets and the criteria for assessing their potential, but on the other hand, it is impossible to ignore the significant downsides of the legislation under this Directive. This downsides especially lie in the strong interdependence between Article 14 and Article 2 of the Directive (EU) no. 2019/1024. Article 2 of the Directive (EU) no. 2019/1024 contains the definitions of the essential terms used in its Article 14, to which the new Section 5b of the Information Act refers. The terms used in Article 14 of the Directive (EU) no. 2019/1024 can be divided into two groups, i.e. technical and institutional terms.

The technical terms including: standard licence, document, reuse, personal data, machine-readable format, open format and high-value data files, among others, appear to be unproblematic because their definitions under Article 2 of the Directive (EU) no. 2019/1024 correspond to the content of the legal order and to the realities of the Czech Republic. However, in relation to institutional terms including: public sector body, bodies governed by public law, public undertaking, university and others, a different conclusion must be drawn. These terms are defined in Article 2 of the Directive (EU) no. 2019/1024 in a way that does not fully correspond to the content of the legal order, to the case law of Czech administrative courts related to the legal order, to the concepts used in the theory of Czech administrative law and to the realities in the Czech Republic or the legal order of the Czech Republic does not even know these terms. As an example, the term university is defined in Article 2 of the Directive as any public sector body that provides post-secondary-school higher education leading to academic degrees (in the Czech Republic is referred to as tertiary education) and the term public sector body is defined in Article 2 of the Directive (EU) no. 2019/1024 as the State, regional or local authorities, body governed by public law or associations formed by

one or more such authorities or one or more such bodies governed by public law (this term is not known in the Czech Republic). Body governed by public law is then defined as the body that has all of the following characteristics: (a) it is established for the specific purpose of meeting needs in the general interest, not having an industrial or commercial character; (b) it has legal personality; and (c) it is financed, for the most part by the State, regional or local authorities, or by other bodies governed by public law; or is a subject to management supervision by those authorities or bodies; or has an administrative, managerial or supervisory board, more than half of whose members are appointed by the State, regional or local authorities, or by other bodies governed by public law (this concept is marginally mentioned in particular by Section 71(3)(b) of the Act no. 235/2004 Coll., on Value-Added Tax, i.e. by a legal regulation in the field of financial law, not administrative law).

Another problem is that Article 14 of the Directive does not even hint at measures leading to the publication and re-use of high-value datasets (which must be compatible with open standard licences), but leaves this issue to subsequent implementing acts, especially to the Implementing Regulation, thus failing to give these measures the appropriate legal force. The effectiveness of the Directive is also undermined by its Article 14(5), which does not sufficiently explain the vague legal term "a substantial impact on the budget of the bodies involved" and it is not sufficiently explained why the requirement to make high-value data files available free of charge does not apply to libraries, including university libraries, museums and archives. The limits of the scope of the Directive (EU) no. 2019/1024 are therefore not entirely clear.

In order to make Section 5b of the Information Act a truly effective piece of legislation, the author *de lege ferenda* proposes the following. Unlike Article 14 of the Directive (EU) no. 2019/1024, the new Section 5b of the Information Act should at least, in principle, define measures leading to the publication and reuse of high-value datasets compatible with open standard licences. Thus, the new Section 5b of the Information Act would, regarding high-value datasets, establish at least some basis for the enforcement of the rights arising therefrom. Furthermore, the new Section 5b of the Information Act should define tools for assessing the potential of high-value datasets, and these tools should correspond to the realities of the Czech Republic. Furthermore, it is essential that the new Section 5b of the Information Act, for reasons of legal certainty, would sufficiently explain any possible exceptions to the requirement to publish high-value data files free of charge. Of course, the most important thing is that the legal regulation of the new Section 5b of the Information Act should be linked, as soon as possible, to an objectively high-quality unique information platform for publishing and searching of high-value datasets. The ideal would probably be to confirm the current trend, with the DIA being the administrator of this information platform, which is already responsible for the administration of the NCOD under Section 4c of the Information Act. The creation of such a platform in the Czech Republic is currently the subject of efforts by the relevant institutions, so we

can only wait to see what the outcome of these efforts will be. The implementation of the *de lege ferenda* proposals of the author would also lead to the fulfilment of the meaning and purpose of Article 288 of the Treaty on the Functioning of the European Union.