



Brussels, **XXX**
[...] (2022) **XXX** draft

COMMISSION IMPLEMENTING REGULATION (EU) .../...

of **XXX**

**laying down a list of specific high-value datasets and the arrangements for their
publication and re-use**

(Text with EEA relevance)

This draft has not been adopted or endorsed by the European Commission. Any views expressed are the preliminary views of the Commission services and may not in any circumstances be regarded as stating an official position of the Commission.

COMMISSION IMPLEMENTING REGULATION (EU) .../...

of XXX

laying down a list of specific high-value datasets and the arrangements for their publication and re-use

(Text with EEA relevance)

THE EUROPEAN COMMISSION,

Having regard to the Treaty on the Functioning of the European Union,

Having regard to Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information¹, and in particular Article 14(1) thereof,

Whereas:

- (1) It transpires from Directive (EU) 2019/1024 that an EU-wide list of datasets with a particular potential to generate socio-economic benefits and with harmonised reuse conditions is a significant enabler of cross-border data applications and services.
- (2) The main objective of establishing the list of high-value datasets is to ensure that public data of highest socio-economic potential are made available for re-use with minimal legal and technical restriction and free of charge.
- (3) Harmonising the implementation of the re-use conditions of high-value datasets entails the technical specification for making the datasets available in a machine-readable format and via application processing interfaces (APIs). Making high-value datasets available under optimal conditions strengthens the open data policies in the Member States, building on the principles of findability, accessibility, interoperability and reusability (FAIR principles).
- (4) Annex I to Directive (EU) 2019/1024 lays down the themes of high-value datasets by listing six thematic data categories: 1) geospatial; 2) earth observation and environment; 3) meteorological; 4) statistics; 5) companies and company ownership; and 6) mobility.
- (5) Following a wide stakeholder consultation and given the result of the impact assessment for this Implementing Regulation, the Commission identified, within each of the six data categories, several datasets of particularly high value and the arrangements for their publication and re-use. The provisions of Union and Member State legislation that go beyond the minimum requirements set out in this Implementing Regulation, in particular in cases of sectoral law, are to continue to apply.
- (6) Pursuant to Directive (EU) 2019/1024, the requirement to make high-value datasets available free of charge shall not apply to libraries, including university libraries, museums and archives. Member States may exempt individual public sector bodies,

¹ OJ L172, 26.6.2019, p. 56.

upon their request and in line with the criteria set out in the Directive from the requirement to make high-value datasets available free of charge for a period not exceeding two years from the date of the entry into force of this Implementing Regulation.

- (7) Directive (EU) 2019/1024 stipulates in its Article 14(3) that this Implementing Regulation 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. However, data held by public undertakings are not included in the scope of this Implementing Regulation.
- (8) Where making high-value datasets available entails the processing of personal data, such processing should be carried out in accordance with Union law on the protection of individuals with regard to the processing of personal data, in particular Regulation (EU) 2016/679 of the European Parliament and of the Council².
- (9) In addition to Directive (EU) 2019/1024, other Union legal acts, including Directive 2007/2/EC of the European Parliament and of the Council³ and Directive 2005/44/EC of the European Parliament and of the Council⁴ may be of relevance for the re-use of public sector information falling within the scope of this Implementing Regulation, notably where those Union acts lay down common requirements for data quality and interoperability.
- (10) Regarding the thematic category ‘companies and company ownership’, Member States are encouraged to go beyond the minimum requirements with respect to the scope of datasets and arrangements for re-use set out in this Regulation.
- (11) Member States should be able to complement the datasets listed in the Annex to this Regulation with public sector information already accessible, whenever such data are thematically related and considered of high value based on the criteria described under in Article 14(2) of Directive (EU) 2019/1024. This includes information that constitutes personal data, whenever Member States consider that this is necessary and genuinely meets objectives of general interest, e.g. anti-money laundering and combatting terrorist financing.
- (12) It is the objective of Directive (EU) 2019/1024 to promote the use of standard public licences available online for re-using public sector information. The Commission’s Guidelines on recommended standard licences, datasets and charging for the re-use of documents⁵ identify Creative Commons (‘CC’) licences as an example of recommended standard public licences. CC licences are developed by a non-profit organisation and have become a leading licensing solution for public sector information, research results and cultural domain material across the world. It is therefore necessary to refer in this Implementing Regulation to the most recent version of the CC licence suite, namely CC 4.0. A licence equivalent to the CC licence suite may include additional arrangements, such as the obligation on the re-user to include

² OJ L119, 4.5.2016, p. 1.

³ Directive 2007/2/EC of the European Parliament and of the Council of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) (OJ L 108, 25.4.2007, p. 1).

⁴ Directive 2005/44/EC of the European Parliament and of the Council of 7 September 2005 on harmonised river information services (RIS) on inland waterways in the Community (OJ L 255, 30.9.2005, p. 152).

⁵ OJ C 240, 24.7.2014, p. 1.

updates provided by the data holder and to specify when the data were last updated, as long as they do not restrict the possibilities for re-using the data.

- (13) The European Data Protection Supervisor was consulted in accordance with Article 42(1) of Regulation (EU) 2018/1725 of the European Parliament and of the Council⁶ and delivered an opinion on [XX XX 2022].
- (14) The measures set out in this Implementing Regulation are in accordance with the opinion of the Committee on open data and the reuse of public sector information set out in Article 16 of Directive (EU) 2019/1024,

HAS ADOPTED THIS REGULATION:

Article 1
Subject matter

- (1) This Implementing Regulation establishes the list of high-value datasets belonging to the thematic categories set out in Annex I to Directive (EU) 2019/1024 and held by public sector bodies among the existing documents to which that Directive applies.
- (2) This Implementing Regulation also lays down the arrangements for publishing and reusing high-value datasets, in particular the applicable conditions for reuse and the minimum requirements for disseminating data via application programming interfaces ('APIs').

Article 2
Definitions

For the purposes of this Implementing Regulation, the following definitions shall apply:

- (1) The definitions set out in Article 2 of Directive (EU) 2019/1024 shall apply.
- (2) The definitions set out for datasets in the geospatial, earth observation and environment categories in Directive 2007/2/EC shall apply.
- (3) The definitions set out for datasets in the mobility category under Directive 2007/2/EC and Directive 2005/44/EC shall apply.
- (4) 'key attribute' means a characteristic of an object or an entity in a dataset, such as a national identification code or name;
- (5) 'granularity' means the level of detail of the dataset;
- (6) 'application programming interface (API)' means a set of functions, procedures, definitions and protocols for machine-to-machine communication and the seamless exchange of data.

⁶ Regulation (EU) 2018/1725 of the European Parliament and of the Council of 23 October 2018 on the protection of natural persons with regard to the processing of personal data by the Union institutions, bodies, offices and agencies and on the free movement of such data, and repealing Regulation (EC) No 45/2001 and Decision No 1247/2002/EC (OJ L 295, 21.11.2018, p. 39).

Article 3

Arrangements for publication applicable to all categories of high-value datasets

- (1) Public sector bodies holding high-value datasets listed in the Annex shall ensure that the datasets described or referenced in the Annex are made available in machine-readable formats via APIs. Where indicated in the Annex, the datasets shall also be made available as a bulk download.
- (2) Those public sector bodies referred to in paragraph 1 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 reuse of high-value datasets laid down in Article 4.
- (3) API terms of use shall be accompanied by API documentation in a Union or internationally recognised open, human-readable and machine-readable format.
- (4) Public sector bodies referred to in paragraph 1 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.
- (5) Public sector bodies holding high-value datasets listed in the Annex shall ensure that the datasets are denoted as high-value datasets in their metadata description.

Article 4

Arrangements for reuse applicable to all categories of high-value datasets

- (1) A decision by a Member State to grant an exemption in accordance with Article 14(5) of Directive (EU) 2019/1024 shall be published online immediately upon its adoption, in the same way as the list of public sector bodies referred to in Article 6(3) of Directive (EU) 2019/1024.
- (2) To facilitate the availability of datasets for re-use covering longer periods of time, the obligations imposed under this Regulation shall also apply to existing high-value datasets created before the date of application of this Regulation.
- (3) High-value datasets shall be made available for reuse under the conditions of the Creative Commons BY 4.0 licence, the Creative Commons Public Domain Dedication (CC0) licence or any equivalent or less restrictive open licence, as set out in the Annex, allowing for unrestricted re-use except for the requirement of attribution, giving the credit to the licensor.
- (4) Public sector bodies holding high-value datasets listed in the Annex may impose additional conditions on the re-use of personal data in the scope of the specific high-value datasets that shall restrict the possibilities for reusing personal data only to the extent necessary to ensure the protection of individuals with regard to the processing of personal data.
- (5) High-value datasets shall be made available in accordance with the arrangements for the publication and reuse as set out in the Annex.

Article 5
Reporting

- (1) By 2 years after entry into force of this Implementing Regulation Member States shall provide the Commission with a report on the measures they have carried out to implement this Implementing Regulation.
- (2) Each Member State shall provide an updated version of the report upon the request of the Commission which should be made every 2 years.
- (3) 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 this 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 reuse of high-value datasets listed in the Annex to this Regulation, per dataset referred to in point b);
 - (c) persistent link to the APIs ensuring access to the high-value datasets listed in the Annex to this Regulation, per dataset referred to in point b);
 - (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 Regulation (EU) 2016/679;
 - (f) the number of public sector bodies exempted in accordance with Article 14(5) of Directive (EU) 2019/1024.

Article 6
Entry into force and application

This Implementing Regulation shall enter into force on the twentieth day following that of its publication in the *Official Journal of the European Union*.

It shall apply from [6 months after entry into force of this Implementing Regulation].

This Regulation shall be binding in its entirety and directly applicable in all Member States.

Done at Brussels,

For the Commission
The President
Ursula VON DER LEYEN