Geospatial Trends 2023

Opportunities for data.europa.eu from emerging trends in the geospatial community

October 2023

data.europa.eu
The official portal for European data

This study has been prepared as part of data.europa.eu. Data.europa.eu is an initiative of the European Commission. The Publications Office of the European Union is responsible for contract management of data.europa.eu.

For more information about this paper, please contact:

European Commission
Directorate-General for Communications Networks, Content and Technology
Unit G.1 Data Policy and Innovation
Email: CNECT-G1@ec.europa.eu

data.europa.eu
Email: info@data.europa.eu

Written by:
Dr Christopher Britsch
Antje Kügeler
Dr Simon Jirka

Last update: 10 October 2023
https://data.europa.eu/

ISBN: 978-92-78-43685-8 doi:10.2830/565958 OA-04-23-813-EN-N

Geospatial Trends 2023

Opportunities for data.europa.eu from emerging trends in the geospatial community

Contents

Executive summary

In this report, current trends regarding geospatial data and their relevance for data.europa.eu are explored. The focus of this annual exercise is to identify opportunities from emerging trends in the geospatial community and to explore how data.europa.eu can support and benefit from the new ways in which geospatial datasets are offered. This year the focus is on geospatial artificial intelligence (GeoAI).

1. Introduction

The goal of data.europa.eu is to improve accessibility and promote the reuse of public sector information. The portal provides access to open data from international, EU, national and regional sources. For the most part, this process is done by collecting the metadata of public data made available across Europe. Metadata is harvested from both specialised geodata and generic open data catalogues and is made available on data.europa.eu as a single point of access.

Geospatial data contains information on properties or conditions that are linked to a position on Earth. Looking at the geospatial context – the 'where' – of a phenomenon will often uncover interesting correlations or revelations. Facilitating the discovery of and access to open geospatial data sources is an important feature of data.europa.eu.

In her article 'Geospatial', Sieber (2019) notes that, while efforts relating to open geospatial infrastructures predate the advent of open data, 'there are relatively weak links between the open geospatial and other open data communities. Stronger links could build critical capacity for spatial analysis within open data communities'. Data.europa.eu offers the opportunity to strengthen the links between open data and geospatial data communities and to promote the use of geospatial information within open data contexts.

As part of the data.europa.eu mission to support European countries in increasing their open data maturity and promoting reuse, the adoption of new trends for publishing geospatial data will be monitored and their relevance for data.europa.eu will be explored on a yearly basis. The focus of this exercise is to identify opportunities from emerging trends in the geospatial community and to explore how data.europa.eu can support and benefit from the new ways in which geospatial datasets are offered. The results are documented in this short report.

2. Trends

In this report, a trend is understood to be a recognisable tendency or direction of development, or, to quote the *Cambridge Dictionary*, a 'general development or change in a situation or in the way that people are behaving' ([1]).

GeoAI – artificial intelligence (AI) for geospatial insights – was chosen to be discussed in this report as it seems especially relevant to data.europa.eu.

2.1. Webinar on geospatial trends

On 8 September 2023, a webinar was held on emerging trends in the geospatial community ([2]). The goal was to gather input from the participants on which trends, standards or ideas could be relevant for data.europa.eu.

At the beginning of the webinar, a brainstorming session on current trends regarding geospatial topics was conducted. The participants were asked what current trends regarding geospatial topics came to mind, and what trends might affect the way we use geospatial data.

([1]) Cambridge Dictionary, 'Trend', 2023, https://dictionary.cambridge.org/dictionary/english/trend.

([2]) Publications Office of the European Union, 'Trends in geospatial data', data.europa.eu website, 27 October 2022, https://data.europa.eu/en/academy/trends-geospatial-data.

Figure 11 depicts a word cloud of the terms entered, where the size of the word is proportional to how frequently it was mentioned by the participants.

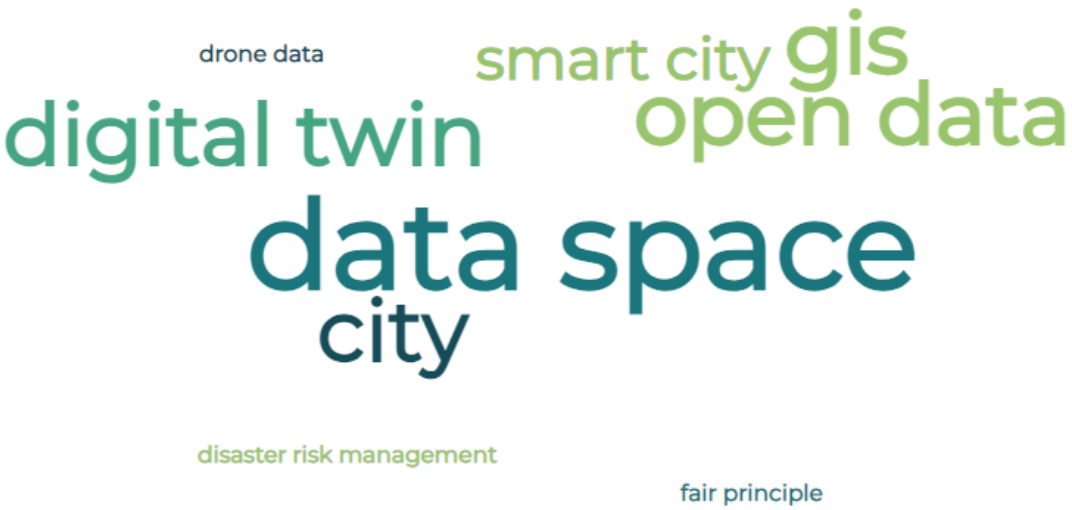


Figure 11: Trends named by participants in webinar on 8 September 2023

While not a representative survey, these insights may be useful in providing input when choosing trends to focus on in next year's report.

This year, the webinar and report focused on GeoAI as a trend that may be of interest for the future of data.europa.eu. After a brief introduction to data.europa.eu, a presentation was given on GeoAI, followed by an online activity to gather participants' views on how data.europa.eu might benefit from and support the use of GeoAI.

The participants responded with many answers centred around aspects such as data and metadata quality. This information is relevant for data.europa.eu to support GeoAI applications in the data-acquisition phase, which is usually a significant challenge. An important goal is to better describe the quality of the data discoverable via data.europa.eu. This goal will require guidance for data providers across Europe to help them better describe the quality of their data so that it becomes usable for GeoAI applications.

Further important suggestions were related to the description of use cases or to use cases in which GeoAI may help to solve societal challenges (e.g. smart applications for citizens, environmental applications, detecting errors in data). To follow this up, a suggestion was to work on success stories and guiding documentation to illustrate how the availability of well-maintained datasets may help to develop GeoAI applications.

Another notable idea by the participants included the organisation of hackathons to foster the use of open data resources for exemplary GeoAI applications. It was suggested to find ways to encourage users to provide feedback to data providers so that they can understand the value of their data for these users. Also, standardisation as a means for simplifying data integration was suggested as an important topic.

2.2. GeoAI

GeoAI can be defined as a machine learning technology that allows the capture and analysis of complex patterns and structures in mainly geospatial data. Machine learning is a subfield of AI in which specialised algorithms are trained with existing data to give predictions for new data. These algorithms often build upon statistical methods and aim to create models from such calculations, which can be used for future predictions. Machine learning models include, for

example, random forest regression, support vector machines or k-nearest neighbour clusters. Other subfields of AI are, for example, robotics or genetic algorithms, which are not part of machine learning or GeoAI. A subfield of machine learning is deep learning, where artificial neural networks – similar to what is done in other machine learning approaches – are trained with existing data to give predictions. These artificial neural networks are highly specialised, algorithm-based digital machines that offer greater learning abilities, robustness and flexibility. Due to recent developments in such algorithms, and in computing power and pretrained neural networks, deep learning is constantly taking up a larger part of machine learning and thus GeoAI approaches.

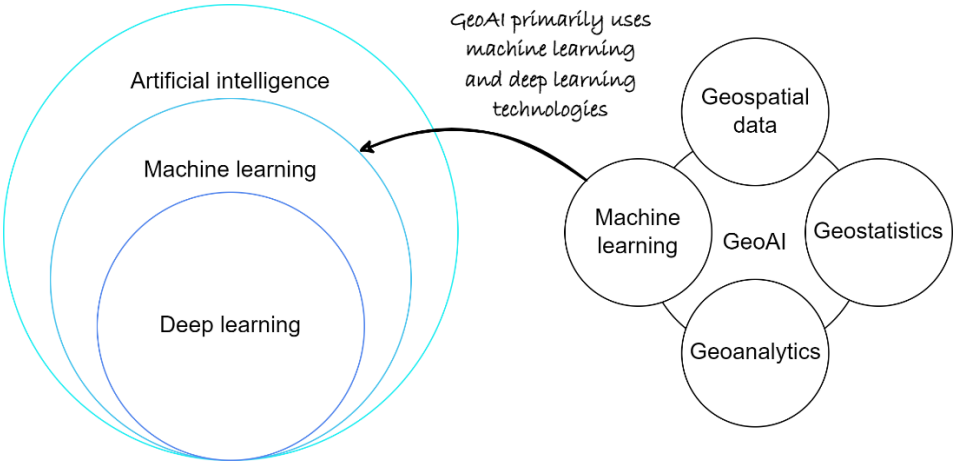


Figure 2: Venn diagram of relations between GeoAI, AI, machine learning and deep learning

The algorithms, models and layers used in GeoAI are not geo-specific, i.e. specific to the geospatial domain. Pretrained models like YOLO, Faster R-CNN or RetinaNet, which are commonly used in any object-detection approach, are used in geospatial analyses as well. Similarly, when constructing individual neural networks from, for example, Keras, TensorFlow or Pytorch layers, the same layers and systematics are used as on any non-geo-specific approach. The 'Geo' in GeoAI derives from the focus on geospatial data. Like in any data-science or machine learning approach, the largest part of a GeoAI project revolves around the data – research, accumulation, migration, transformation, preparation, preprocessing – which uses up to 80 % of any given GeoAI project's time and effort. To perform these steps on geospatial data, specialised knowledge on said data types and formats is required. This data allows users to analyse different spectral layers of satellite images, predict local or regional events or sort out complex errors in large data masses. Why, however, is data so relevant for AI and taking up such a large part of an AI project?

The whole idea behind a machine-learning application (a subpart of AI, including deep learning with artificial neural networks) is the training of an algorithm on existing (past) data to give a prediction on new (current or future) data. Therefore, data is, in a way, the engine and fuel for any machine-learning approach. Without an adequate amount of data, no algorithm will be able to learn enough interrelations in the data to give any predictions. Likewise, the quality of the data determines the quality of the information output.

This need puts a large spotlight on open data. The accessibility of data – especially when available for different regions – can easily be the deciding factor of success for any AI approach. This data accessibility is especially relevant when regarding projects addressing environmental or climate issues. To analyse which preventive (often invasive) action in nature is necessary and justified – for example the use of pesticides or felling of trees – statistical methods are used for calculations. However, due to dynamic climate changes, weather-dependent systems, for example, are not well observable anymore through common statistical methods. Artificial neural networks offer a greater robustness and flexibility in that matter. Therefore, GeoAI represents a great power in supporting preventive action.

Nonetheless, any GeoAI model is only as good as its input data. Wide range accessibility of data does not only ensure the necessary quantity, but also allows a broader analysis and comparison of parameters, ultimately allowing a better judgement of quality.

In 2018, the European Commission published the short paper 'AI and open data: A crucial combination', which highlights the relevance of volume, variety and veracity of input data for AI approaches. The publication emphasises the importance of availability of high-quality data in high volume to achieve successful AI projects. It further states the relevance of variety in geographic origin of data, and that 'data is often geographically or culturally biased'. While referring, in this case, to data from consumer surveys, this is just as – if not even more so – true for any GeoAI approach. Logically, the geographical aspect has a lot of influence in any geospatial study. This influence ranges from a diversification of data regarding, for example, weather phenomena to geological attributes of different areas. Information from satellite imagery differs greatly depending on surfaces. Weather influences behave differently in different regions. Plant and animal life show differences and varying behaviours. Human influences in traffic, building structures and use of space, among other things, project differently. But aside from these, often visual, effects directly traceable in the data, there is another important factor: the individuality of handling data (which is also addressed in the European Commission's paper). The data, its quality and the extent to which it is recorded and published can be greatly dependent on the country, municipality or institution from which it originates. Such diversity can be an advantage. However, it can far more often be a challenge for performing AI methods up to the point of rendering them impossible. Too much diversity in quality, completeness, accuracy or up to dateness can make it impossible to employ AI methods, since any AI algorithm needs to be able to find rules within a system. The impressive aspect in any AI approach is the possibility to analyse a system the rules of which we are not aware. This is simply due to the fact that often there are underlying rules that cannot be observed by humans, because they exist, for example, in a multidimensional space. However, as good as AI algorithms are at finding such rules, they still need relatively homogeneous data to allow them to discover these rules. To put it in exaggerated terms, if every data point in a dataset is collected in a different way, the exception is the rule and thus no interrelations can be found, and no predictions can be made.

This requirement elevates the relevance of centralised data portals like data.europa.eu, containing detailed information on data from reliable sources. Current and especially future problems can and will be solved to a large extent by employing AI models. Regardless of whether addressing climate issues or complex infrastructures, capturing the intricacy of our evolving, changing and adapting world will need the help of high-performance algorithms. These algorithms, as explained above, will need extensive high-quality data from as many sources as possible in a high volume. More importantly, this data needs to be easily accessible and comparable. By diminishing borders and ensuring relevant standards, data.europa.eu can make a great contribution to the advances in using GeoAI for important and necessary purposes. Through its centralised position, data.europa.eu can not only ensure accessibility, but also propose standards for high quality and comparability to ultimately fuel future developments of GeoAI projects.

3. *Key points for attention from the 2022 report*

As a result of our analysis from 2022, we have listed some potential work items that would be useful in further advancing data.europa.eu. These key points for attention are described in this section.

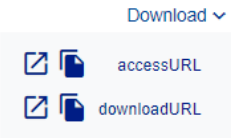
3.1. Promote OGC API Features endpoints and their homogeneous description in geocatalogues using International Organization for Standardization metadata

The 2022 report named some possible steps to be considered in 2023. The investigation is summarised in Table 1.

Table 1: Promote Open Geospatial Consortium Application Programming Interface (OGC API) Features endpoints and their homogeneous description in geocatalogues using International Organization for Standardization (ISO) metadata

	Topics to investigate	Results
1	Are there any efforts to harmonise the way OGC API Features are described in ISO geo-metadata or 'geo' Data Catalogue Vocabulary Application Profile (DCAT-AP)?	There seem to be no ongoing plans for harmonisation. However, some helpful answers to our question were supplied in the Infrastructure for Spatial Information in Europe (Inspire) discussion forum from a team member for the operational support to the maintenance and technical evolution of Inspire components and artefacts ([3]) (see topic 2).

2	How can OGC API Features endpoints described in geo-metadata be identified in the geo-harvesting process?	According to the 'Data service linking simplification good practice' ([4]) dataset, metadata should be used. The use of the gmx:Anchor encoding is recommended, but the <gco:CharacterString> can also be used.
		
		The following are examples of the different encodings:
		
		Recommended:
		
		`<gmd:protocol>`
		`<gmx:Anchor xlink:href="http://www.opengis.net/def/docs/17-069r3"> OGC API Features`
		`</gmx:Anchor>`
		`</gmd:protocol>`
		
		Also permitted:
		
		`<gmd:protocol>`
		`<gco:CharacterString>OGC API Features`
		`</gco:CharacterString>`
		`</gmd:protocol>`

3	How can OGC API Features endpoints be displayed as a distribution?	A distribution link as either accessURL or downloadURL is appropriate. 

([3]) GitHub, 'Describing OGC API Features in ISO geo-metadata #161', INSPIRE-MIF website, 28 August 2023, https://github.com/INSPIRE-MIF/helpdesk/discussions/161.

([4]) GitHub, 'Data service linking simplification: Good practice guidelines', INSPIRE-MIF website, 20 November 2022, https://github.com/INSPIRE-MIF/gp-data-service-linking-simplification/blob/main/good-practice/data-service-linking-simplification-spec.md#811-inspire-data-set-metadata-resource-locator–requirements-class-.

4	What would be a user-friendly way to filter for data provided via OGC API Features on data.europa.eu?	The filter for 'format' includes other types of geo-services, such as the Web Map Service or the Web Feature Service. These are not strictly 'formats' but ways to access the data, as the geo-services can often deliver different data formats (e.g. Geography Markup Language, GeoJSON, Keyhole Markup Language). Following this example, it would be good to add 'OGC API Features' as a 'format'.
5	After a way to identify metadata on OGC API Features endpoints has been established, how can a first version of an enhanced version of the GeoViewer be designed?	It is possible to identify OGC API Features endpoints in ISO metadata if the approach described in topic 2 above is applied by the majority of data providers. The metadata should include a URL (uniform resource locator) to the landing page. The landing page for OGC API Features is the top-level resource that serves as an entry point ([5]). From the landing page, a client application can retrieve the URLs of the collection and API definition paths that are needed to access the data to display it in an application like the GeoViewer. Should the GeoViewer be enhanced to allow a preview of OGC API Features, further business logic would be needed to support the selection of data from specific collections and to support the paging mechanism offered by OGC API Features endpoints to handle larger feature collections.

3.2. Investigate the inclusion of SensorThings API instances and how they can be represented as dedicated service types on data.europa.eu

The 2022 report named some possible steps to be considered in 2023. The investigation is summarised in Table 2.

Table 2: Investigate the inclusion of SensorThings API instances and how they can be represented as dedicated service types on data.europa.eu

	Topics to investigate	**Results**
1	How will content offered by SensorThings API instances be made discoverable via data.europa.eu?	In general, the concept of data streams offered by the SensorThings API standard allows users to organise the provided sensor data into the equivalents of datasets. Thus, we recommend enabling the discovery of SensorThings API contents based on data streams.
2	Will it be possible to discover individual data streams, or is it better to just discover the server instances as a whole? Are there hybrid approaches that offer a suitable trade-off between both ideas?	An analysis of SensorThings API endpoints available to the public has shown that very different granularities were implemented. Although the general recommendation is to enable the discovery of data streams (see topic 1 above), certain SensorThings API servers offer a very high number of data streams representing the data delivered by large numbers of individual devices. For these cases, an aggregated discovery of the data offered by a SensorThings API server would be advisable. As a result, we recommend to further investigate hybrid approaches so that the granularity of the metadata and discovery can be adjusted to the actual number of data streams offered by a server.

([5]) Open Geospatial Consortium, 'Resources of OGC API – Features', OGC e-Learning website, 6 September 2023, http://opengeospatial.github.io/e-learning/ogcapi-features/text/operations.html.

3	How can the metadata-harvesting process be designed in an automated way to keep the workload for human administrators as low as possible?	The SensorThings API does not offer a landing page that could be used for deriving metadata (unlike, for example, the OGC API Features). However, based on the SensorThings API resource types (DataStream, Sensor, Location, ObservedProperty), an automated harvesting of metadata is possible. Basically, these resource types allow a metadata harvester to determine which data is offered by a SensorThings API instance. However, as data.europa.eu does not interact with the SensorThings API instances directly but harvests metadata from catalogues, we recommend reaching out to interested stakeholders to investigate good practices on how to provide metadata for SensorThings API instances.
4	How can the metadata describing a SensorThings API instance be mapped to DCAT-AP?	There are already first descriptions of SensorThings API instances available in different catalogues. However, a major challenge remains, which is to further investigate the optimal level of granularity to use on the representation of the available data (see topic 2).

3.3. Advance the geo-visualisation component to support the OGC SensorThings API standards as a further source of real-time data

The 2022 report listed some possible steps to be considered in 2023. The investigation is summarised in Table 3.

Table 3: Advance the geo-visualisation component to support the OGC SensorThings API standards as a further source of real-time data

	Topics to investigate	**Results**
1	How can the preview of SensorThings API instances be supported in the geospatial data visualisation?	There are different ways of how to enable the preview of SensorThings API data. Important challenges that need to be addressed are the provision of sufficient metadata via the harvested national catalogues and deciding upon which type of resources shall be visualised (see topic 2 below). Based on this, we recommend initiating development activities to investigate how the connection of selected SensorThings API instances to the geospatial data visualisation component of data.europa.eu could be implemented (on a prototype level).

2	What data should be visualised (e.g. sensor locations, latest measurements, time series of historical measurements)?	The sensor locations would be the initial information that should be visualised. The visualisation of further types of information strongly depends on the kind of data offered by the SensorThings API instances. In the case of measurements of basic properties (e.g. simple scalar values), the previews of the latest measurement would also be useful. Other types of data (e.g. location observations of moving objects) would require different types of visualisation. Thus, further investigation on recommendations for the preview of different types of data would be a very useful input. The preview functionality of the geospatial data viewer is not intended to become a comprehensive tool for sensor data exploration. However, depending on the type of data offered by a SensorThings API server, the preview of time series data as overview diagrams could be a useful functionality.
3	Are conceptual changes necessary to accommodate the preview of the SensorThings API (e.g. new workflows for data selection)?	Yes, the SensorThings API model would necessitate some enhancements of the geospatial data viewer. This comprises, on the one hand, more comprehensive data selection functionality (e.g. for selecting data streams). On the other hand, new types of data, such as time series or live location information, will also require new approaches for data visualisation. To start with this task, the work mentioned under Section 3.2 needs to be completed first.

3.4. Investigate ways to support finding and accessing high-value datasets

The 2022 report suggested that methods be investigated to support finding and accessing high-value datasets (HVDs).

The concept of HVDs was introduced in Directive (EU) 2019/1024 (the open data directive). The objective of this 2019 recasting of the public sector information directive was to strengthen the EU's data economy by increasing the amount of publicly funded data available for reuse. HVDs are datasets meant to be used to create value-added services to benefit society, the environment and the economy. HVDs must be made available for reuse:

- with minimal legal restrictions,
- free of charge,
- in a machine-readable format,
- as a bulk download (where relevant),
- via suitable APIs.

On 20 January 2023, the European Commission passed an implementing regulation laying down a list of specific HVDs and the arrangements for their publication and reuse. Member States are required to make these datasets available within 2 years from the date of the entry into force.

Table 4: Investigate ways to support finding and accessing high-value datasets

	Topics to investigate	Results
1	How do the Member States publish and describe HVDs?	The discussion on how to publish and describe HVDs is ongoing ([6]). Thus, we recommend following the current discussions to derive recommendations for data.europa.eu.
2	What insights does the planned new feature of the Inspire geoportal ([7]) for browsing/viewing geospatial HVDs offer to users? How many datasets can be discovered in this way? Can this approach help users locate geospatial HVDs from authoritative governmental sources?	In its annexes, the Inspire directive lists 34 data themes that fall within its scope. Annex I to the open data directive lists six thematic categories. The implementing regulation for the HVDs references Inspire annex themes in three of the six thematic categories. This mapping between Inspire and HVD themes will be used for the new Inspire geoportal to show which datasets are HVDs. As this functionality is not yet available, it cannot be determined how many datasets can be discovered in this way. It is expected that this approach will help users locate geospatial HVDs from authoritative governmental sources, but there is concern that there will be a very large number of datasets, which may make it difficult to identify the datasets needed for a specific use case.
3	What are users' expectations for data.europa.eu in regard to finding and accessing HVDs? Are there specific user expectations for data.europa.eu in regard to finding and accessing geospatial HVDs?	These questions could be posed in an online user survey. We recommend conducting a user survey on this or including these questions in the next user survey.
4	Is there a need for data.europa.eu academy training in regard to HVDs, for example explaining the relationship between Inspire annex themes and HVDs?	As this information is difficult to find and to understand if you are not an expert, this could be a good topic for additional training sessions.

3.5. Investigate how metadata from OGC API Records endpoints can be harvested efficiently

At the time of writing this report, the OGC API Records specification ([8]) was still in draft form. A currently empty repository ([9]) has been created as a discussion space on possible good practices for Inspire catalogue services based on OGC API Records.

As the investigation should be carried out after the OGC API Records specification has been published, we will put this on the agenda for 2024.

([6]) Fragkou, P. (2023), *DCAT-AP High Value Datasets*, https://semiceu.github.io/DCAT-AP/releases/2.2.0-hvd/.
([7]) European Commission, 'Inspire geoportal', European Commission website, https://inspire-geoportal.ec.europa.eu/.
([8]) Open Geospatial Consortium, 'OGC API – Records', Open Geospatial Consortium website, https://ogcapi.ogc.org/records/.
([9]) GitHub, 'Discussion space on a possible good practice for Inspire catalogue services based on OGC API – Records', INSPIRE-MIF website, 8 November 2021, https://github.com/INSPIRE-MIF/gp-ogc-api-records.

Table 5: Investigate how metadata from OGC API Records endpoints can be harvested efficiently

	Topics to investigate	Results
1	When will the first OGC API Records catalogues be available for harvesting?	Unclear, as the OGC API Records specification is still a draft. This should be revisited once the specification is finalised.
2	Will Inspire recommend that Member States switch from Catalogue Service for the Web to OGC API Records?	Unclear, as the OGC API Records specification is still a draft. This should be revisited once the specification is finalised.
3	Are there benefits in adding OGC API Records to the endpoints supported by the geo-harvester?	Unclear, as the OGC API Records specification is still a draft. This should be revisited once the specification is finalised.

3.6. Investigate implications from smart city initiatives

The 2022 report suggested to explore implications from smart city initiatives for data.europa.eu. The results of this exercise are summarised in Table 6.

Table 6: Investigate implications from smart city initiatives

	Topics to investigate	Results
1	What are the relevant types of data sources offering smart city data?	There is a very broad range of different data source types that offer smart city data. Typical approaches comprise the OGC SensorThings API, the FIWARE software framework, as well as technologies related to the internet of things (e.g. the Message Queuing Telemetry Transport (MQTT) protocol).
2	What are suitable ways to provide metadata about smart city data resources?	Generally, DCAT-AP is a suitable metadata format and there are already several smart city resources discoverable via data.europa.eu. However, further guidance for data providers will be needed on how to describe additional types of relevant data sources using DCAT-AP so that they can be easily used from client applications. Thus, we recommend initiating an additional activity as part of data.europa.eu to develop guidance for data providers on how to describe typical smart city resources (e.g. SensorThings API servers, MQTT-based data streams (see topic 1)) via DCAT-AP.
3	How can smart city initiatives be motivated to share their data as open data and to provide sufficient metadata to make it discoverable?	To promote the sharing of smart city data, activities on different levels are recommended. On a political level, incentives for the sharing of smart city data should be developed. On a practical level, the provision of success stories to illustrate the benefits of data sharing would be helpful to serve as examples to motivate further data providers. Finally, the provision of technical guidance (see topic 2) and the provision of corresponding tooling (e.g. to automate the generation of metadata files) are also recommended.

4. *Conclusions and key points*

Geospatial data can play a key role in addressing global challenges such as environmental pollution, economic crisis, loss of biodiversity and climate change. New technological means change the way we work with data. This change may lead to new opportunities and ideas to improve data.europa.eu.

As a result of our analysis, we see that the potential work items listed below would be particularly useful in further advancing data.europa.eu.

This report has been compiled to investigate and start a discussion on how current geospatial trends might affect user expectations for data.europa.eu and to find opportunities that arise from them. The ideas from the *Geospatial Trends 2022* report ([10]) were discussed against the background of current developments. Some of them remain open to be revisited again, for example when the relevant legislation has been passed or standards have been adopted. In addition to this, new key points of attention to follow up on over the next years were identified. These will be addressed in next year's report on geospatial trends.

4.1. Investigate the potential role of data.europa.eu in current GeoAI developments

Topics to investigate are listed in Table 7. They should be addressed in 2024.

Table 7: Investigate the potential role of data.europa.eu in current GeoAI developments

	Topics to investigate
1	In what way can data.europa.eu support the use of GeoAI?
2	How can we create synergies between data.europa.eu and research data infrastructure projects, e.g. the European Open Science Cloud ([11])?
3	How can we develop a list of data quality properties that are needed by thematic experts to discover and assess the suitability of using specific datasets for GeoAI applications?
4	How can we develop recommendations on how the necessary information about data quality and contents of datasets can be encoded efficiently in DCAT-AP? Based on this information, data scientists will become better able to use data.europa.eu for discovering and evaluating datasets for their GeoAI applications.

([10]) Kügeler, A. and Jirka, S. (2022), *Geospatial Trends 2022 – Opportunities for data.europa.eu from emerging trends in the geospatial community*, Publications Office of the European Union, Luxembourg, https://data.europa.eu/sites/default/files/course/Geospatial_Trends_2022.pdf.

([11]) European Open Science Cloud, 'The European Open Science Cloud Portal', European Open Science Cloud website, https://eosc-portal.eu/.

4.2. Investigate ways to support finding and accessing high-value datasets

As the discussion on how to publish and describe HVDs is still ongoing, it will be useful to revisit the questions from this year's report in 2024. These topics are listed in Table 8.

Table 8: Investigate ways to support finding and accessing high-value datasets

	Topics to investigate
1	How do the Member States publish and describe HVDs?
2	What insights does the planned new feature of the Inspire geoportal ([12]) for browsing/viewing geospatial HVDs offer to users? How many datasets can be discovered in this way? Can this approach help users locate geospatial HVDs from authoritative governmental sources?
3	What are users' expectations for data.europa.eu in regard to finding and accessing HVDs? Are there specific user expectations for data.europa.eu in regard to finding and accessing geospatial HVDs?
5	Is there a need for data.europa.eu academy training in regard to HVDs, for example explaining the relationship between Inspire annex themes and HVDs?

4.3. Investigate how metadata from OGC API Records endpoints can be harvested efficiently

The following tasks can be addressed after the specification for OGC API Records has been published (it is currently in draft form). Topics to investigate are listed in Table 9.

Table 9: Investigate how metadata from OGC API Records endpoints can be harvested efficiently

	Topics to investigate
1	When will the first OGC API Records catalogues be available for harvesting?
2	Will Inspire recommend that Member States switch from Catalogue Service for the Web to OGC API Records?
3	Are there benefits in adding OGC API Records to the endpoints supported by the geo-harvester?

([12]) European Commission, 'Inspire Geoportal', European Commission website, https://inspire-geoportal.ec.europa.eu/.

5. Bibliography

European Commission (2018), 'AI and open data: A crucial combination', data.europa.eu website, https://data.europa.eu/en/publications/datastories/ai-and-open-data-crucial-combination.

European Commission, Commission Implementing Regulation (EU) 2023/138 of 21 December 2022 laying down a list of specific high-value datasets and the arrangements for their publication and re-use (Text with EEA relevance), http://data.europa.eu/eli/reg_impl/2023/138/oj.

European Commission (2022), Commission staff working document – Evaluation of Directive 2007/2/EC establishing an infrastructure for spatial information in the European Community (INSPIRE), SWD(2022) 195, https://ec.europa.eu/info/law/better-regulation/have-your-say/initiatives/12427-Sharing-geospatial-data-on-the-environment-evaluation-INSPIRE-Directive-en.

European Commission (2022), Commission staff working document – Executive summary of the evaluation of Directive 2007/2/EC establishing an Infrastructure for Spatial Information in the European Community (INSPIRE), SWD(2022) 196, https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52022SC0196.

European Parliament and Council of the European Union, Directive 2003/4/EC of the European Parliament and of the Council of 28 January 2003 on public access to environmental information and repealing Council Directive 90/313/EEC, OJ L 41, 14.2.2003, p. 26, http://data.europa.eu/eli/dir/2003/4/oj.

European Parliament and Council of the European Union, 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, http://data.europa.eu/eli/dir/2019/1024/oj.

European Parliament and Council of the European Union, Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the reuse of public sector information (open data directive), OJ L 172, 26.6.2019, p. 56, http://data.europa.eu/eli/dir/2007/2/oj.

Jirka, S. and Kügeler, A. (2022), *Real-time Data 2022: Approaches to integrating real-time data sources in data.europa.eu*, Publications Office of the European Union, Luxembourg, http://data.europa.eu/doi/10.2830/182607.

Kügeler, A. and Jirka, S. (2022), *Geospatial Trends 2022 – Opportunities for data.europa.eu from emerging trends in the geospatial community*, Publications Office of the European Union, Luxembourg, http://data.europa.eu/doi/10.2830/077789.

Kügeler, A. and Jirka, S. (2021), *Geospatial Trends 2021: Opportunities for data.europa.eu from emerging trends in the geospatial community*, Publications Office of the European Union, Luxembourg, https://data.europa.eu/sites/default/files/report/Geospatial_Trends_2021_1.pdf.

Sieber, R. (2019), 'Geospatial', in Davies, T., Walker, S. B., Rubinstein, M. and Perini, F. (eds), *The State of Open Data – Histories and horizons*, African Minds, Cape Town, pp. 137–150 https://library.oapen.org/handle/20.500.12657/24884.

ISBN: 978-92-78-43685-8

Publications Office
of the European Union