

Checklist for sharing data

This is a pdf summary of the operating model for data sharing. Find out more about the Operating model for data sharing from the Open Data service at: www.avoindata.fi/en/operating-model

1. Familiarise yourself with the backgrounds and ecosystems of data sharing

- Familiarise yourself with information policy and the sharing of data, as well as the significance of knowledge-based management.
- Make sure that the people involved in sharing data in your organisation have sufficient understanding and competence.
- Provide training if necessary.
- Familiarise yourself with open-data ecosystems and different cooperation networks.

2. Identify any restrictions or obligations

An organisation that is considering sharing its data must carefully familiarise itself with the different obligations and legislation that apply to its field. For example, see:

- Act on Information Management in Public Administration
- Act on the Openness of Government Activities
- Data Governance Act
- General Data Protection Regulation and Data Protection Act
- Archives Act
- Open Data Directive
- If your organisation produces spatial data, see the INSPIRE Directive

If necessary, consult with your organisation's legal experts and data protection officer about the data sharing process.

3. Plan your data sharing objectives; organise and allocate resources

Take care of your organisation's information management. Make sure that data sharing is included in the organisation's goals and taken into account in, for example, its operational and financial plans, to ensure continuity and sufficient resources. The organisation's commitment to opening its data can be strengthened with various incentives and requirements.

Ensure that your organisation's information management is organised in accordance with the obligations laid down in the Public Information Management Act. Pay particular attention to these factors:

- The tasks of an information management entity
- Responsibilities of an information management entity's management body

When allocating resources to data sharing, pay attention to e.g. the following costs:

- Human resources

- Dataset processing
- Dataset storage
- Dataset descriptions (metadata)
- Dataset publication and communications
- Dataset maintenance and updates

Organise the process for opening the data:

- Define responsibilities, for example
 - Person(s) responsible for sharing data
 - Information security and data protection experts
 - Technical experts for planning the technical implementation of data sharing and processing
 - Person responsible for communications
- Plan the process

Develop the process of opening data openly, in cooperation with stakeholders.

4. Survey what data is suitable for sharing and the demand for it

When surveying the organisation's data, take into account the potential social, environmental, and economic value of different datasets. If possible, prioritise the opening of datasets that promote equality, knowledge-based decision-making or the circular economy.

The surveying process can make use of e.g.:

- The organisation's information management model
- The public administration information management map
- The organisation's document publicity description
- The organisation's data balance sheet
- Interoperability tools

The opening of certain datasets is subject to specific legal provisions, which is why their opening should be prioritised first. These include the following **categories of high-value datasets**:

- Geospatial
- Earth observation and environment
- Meteorological
- Statistics
- Company and company ownership
- Mobility

The opening of data should be based on demand – if the data is not used, it will not create any value. When you know what type of data your organisation has, determine the data for which there would be demand outside your organisation and what benefits could be gained from opening the data.

When surveying the demand for data, make use of the following:

- Information requests and feedback received by the organisation
 - If certain datasets are requested frequently, it is advisable to make them openly available to everyone, if possible.

- User statistics of the organisation's website
 - If a topic attracts interest on the organisation's website, there may be demand for open data related to it.
- User surveys
 - Potential data users can be asked about the type of data they would find the most useful.

5. Evaluate and define the properties of the data to be shared

Once you have identified the data that is suitable for sharing, you will need to determine

- who manages and is responsible for the data and its underlying information system,
- whether there are national or international standards for sharing this data, and
- whether another party has already opened any similar data, in which case its data model could be utilised in the opening of your data.

You will also need to assess the following factors related to the data to be shared:

- whether it is subject to e.g. any copyright and information security factors that may limit its sharing,
- the benefits, risks, and costs of sharing, e.g. through the BRC evaluation tool,
- its quality, e.g. with data quality criteria, and
- the need for anonymisation and aggregation, by examining whether the data is public or if it contains personal data or other information critical to the functioning of society.

Define the following for the shared data:

- the most suitable form of distribution, and
- the licence under which the data will be shared.

Decide on the sharing of the data with the organisation's management, if necessary, and manage the residual risks associated with data sharing, i.e. risks that remain in force or cannot be addressed or are best left unaddressed.

6. Describe and publish the data and communicate about it

Describe the data to be published comprehensively. The data published in the data portals contains basic descriptive information, such as the dataset's name and licence. In addition to basic information, the data's metadata should also describe:

- the data's content,
- the data's generation process,
- the data's quality, and
- the data's administrator.

If possible, any Finnish data should also be described in Swedish and English, to facilitate its international use.

If you use e.g. the Open Data service in the publication of the data, the data must be described in accordance with the **DCAT-AP data model**.

Publish the data and its metadata

- Ensure that the data to be published complies with the **FAIR principles**.
- Publish the data in a suitable data portal, such as the public **Suomi.fi Open Data** portal, where the data can be found easily and quickly.

Communicate about the data's publication

- Inform potential users of the publication, for example on social media and in the organisation's newsletter.
- Explain what type of data has been published and where.

7. Maintain and update the published data

Strive to continuously improve the data's quality, and develop the dataset based on the feedback you receive.

Update the opened data

- Make sure that the dataset is updated according to the frequency specified in the metadata.
- In addition, make any older versions of the dataset available when possible.

Update the metadata

- Please note that it is important to keep the metadata of the dataset up to date and, if necessary, update it at the same time as the dataset is updated.
- For example, if the data has been published through a continuously updated API, update the metadata when a new version of the API is published or any changes are made to its other data.

Inform potential users of any changes to the published data or API in the suitable communication channels.

Encourage users to report any errors they detect and develop the data based on the feedback. Plan in advance how your organisation will react if an error is detected in the published data.

Depending on the severity of the error:

- The dataset can be completely removed from distribution until further notice.
- If the error is not critical, mention it in the dataset's metadata and describe any corrections.
- Inform the potential users of the data about the error in the suitable communication channels.

8. Monitor the use and benefits of published data

The use of data and its impact can be monitored by:

- collecting statistics on data downloads and, if possible, use cases.
- monitoring data use in new services and applications as well as on the media.
- conducting a cost-benefit analysis on the economic benefits of opening the data. However, please note that the monetary value of many benefits can be difficult to measure.
- commissioning reports or studies.

When assessing the impacts, it is advisable to take into consideration the entire life cycle of the data and the impacts of opening it from a wider societal perspective.

9. Support data users and collect feedback

The use of shared data should be supported using other means, such as effective communications, as the mere act of opening data does not create value in itself.

The interaction between the data's users and your organisation increases the utilisation of your published data and can also help to improve the data's quality.

- **Offer different feedback channels**, such as email and social media channels, through which data users can easily send feedback and development suggestions.
- **Work together** with developers and organise activities around the data, such as developer meetups and hackathons.

10. Leave the data available even if it is no longer updated

Datasets published as open data should be publicly available for as long as possible.

If you cease updating a dataset:

- Do not remove it from distribution unless it is absolutely necessary.
- State in the metadata why and when updates were discontinued.
- Please note that even if the dataset is removed, its users may still continue to use the dataset in accordance with its licence.

If you intend to discontinue an API:

- Communicate about its discontinuation well in advance, as the API may be used in other services.