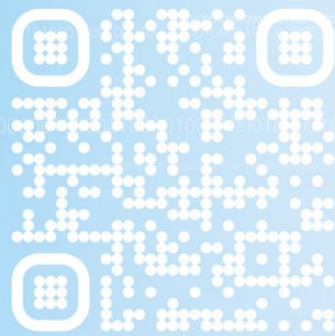


# A Guideline for Open Data

Bernhard Krabina

data.  
gv.  
at



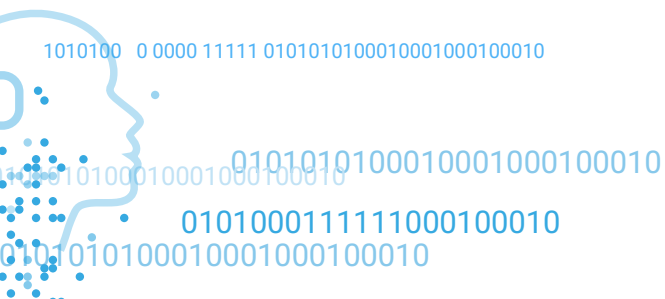
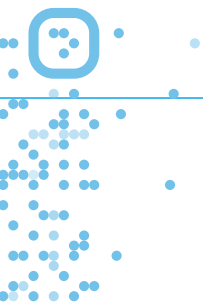
**Interreg**  
CENTRAL EUROPE



Co-funded by  
the European Union

DoorCE

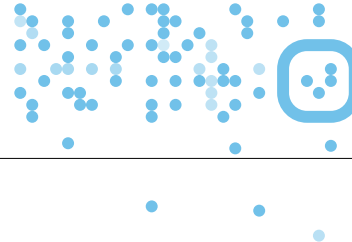
OPEN DATA ■ BENEFITS ■ ALL



---

# Contents

Preface.....	4
The Essentials on One Page .....	5
How to Proceed? – Steps to Open Data .....	6
Step 1: Open Licenses (Legal Aspect).....	7
Step 2: Machine-Readable Formats (Technical Aspect).....	8
Step 3: Metadata (Organizational Aspect) .....	8
Which Data? – Data Catalog.....	9
Guidelines for Open Data .....	10
Criteria for Data Monitoring.....	11
What to Consider? – The Legal Framework .....	12
Access to Information .....	12
Framework Conditions for Further Use .....	13
Open by Default.....	14
Where to Publish? – Data Portals .....	15
How to Start? – Phases of Implementation .....	16
Why Open Data? – Open Government.....	18
What to Expect? – Frequently Asked Questions / FAQ .....	20
Which Terms Should You Know? – Glossary .....	23
Where Can You Read More? – Further Information.....	24
Guides and Studies.....	24
Internet Links .....	24
Bibliography .....	25
Imprint.....	26



# Preface

Digital data has become indispensable in our daily lives. Open Data is the term used for data that is made available to the public free of charge for further use. The public sector plays a central role in this: Authorities generate and publish data, act as regulators and also provide data infrastructures. In Austria, data.gv.at has been available for over 10 years, a free infrastructure for all authorities to publish freely available data. The Cooperation OGD Austria connects the federal government, states, cities and municipalities with experts from science and research, develops guidelines and standards and networks stakeholders internationally.

Many national and international issues present society with diverse challenges, for which broadly available and usable data is an essential basis.

Open data is the cornerstone of open knowledge for our society. So that we can jointly make evidence-based decisions for the common good.

The guide to open data aims to summarize the essentials on the topic for employees in the public sector and to motivate them to further engage with the topic. With the guide, all public bodies are also well prepared for current and future legal obligations for the publication of information.

**Brigitte Lutz**

Spokesperson Cooperation OGD Austria, Data-Governance Coordinator of the City of Vienna

**Brigitte Barotanyi**

Federal Ministry of Finance, Department V/6, E-Government Federal Government/Administration

**Robert Seyfriedsberger** und **Michael Reichart**

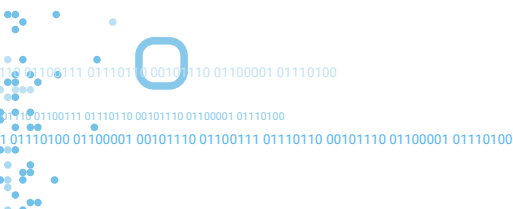
Federal Computing Center, Open Data Team

**Bernhard Krabina**

Open Data Expert, IKT Linz



*An example of the importance of open data is the data material on the COVID-19 pandemic: Since the authorities made the data basis available for free further use, over 85 different applications and visualizations have been developed in Austria.*



# The Essentials on One Page

Numerous public bodies have recognized that, in addition to offering information on the internet in the form of websites or documents (PDF), providing data in a reusable form is a contemporary way of publishing.

Open administrative data (Open Government “Data”, OGD for short) are therefore not personal data that is collected, created or paid for by public bodies and made available to the public free of charge.

Open administrative data constitute a basic infrastructure for digitization. The aim is to make the data as easy as possible to reuse: by business, science, media and civil society and to create new applications (apps) or visualizations from it or to contribute to information and transparency.

Many authorities in Austria already publish data in a reusable form and list it on the Austrian data portal [data.gv.at](https://data.gv.at). This is operated by the Federal Computing Center (BRZ) on behalf of the Federal Ministry of Finance (BMF) and is available free of charge to all authorities in Austria. The “[Cooperation OGD Österreich](#)” networks administrations and experts and develops framework conditions for OGD in Austria.

What are data anyway? Technically speaking, PDF documents or the internet pages of your authority are also data. However, open data refers in particular to structured data, i.e. e.g. data that is printed in the form of a table in a PDF document, or the underlying data that is visualized using a diagram or on a map.

The publication of open data is essentially carried out in three steps:

1. [under the specification of a free license](#), that specifies how users may use the data (legal aspect),
2. [in open and machine-readable technical formats](#) (technical aspect),
3. [described by metadata](#) (information about the published data, organizational aspect).

The guide to open data contains further information on data catalogs, data portals and Open Government in addition to the description of these most important steps. Different phases of implementation as well as questions and answers are also described..

If you have never heard of DCAT-AP, JSON or APIs, you have come to the right place! This guide aims to convey the most important things to you on about 20 pages and to refer to numerous further sources.



# How to Proceed? – Steps to Open Data





Open data is a complex topic. But actually, only three essential steps are necessary to open data..

## Step 1: Open Licenses (Legal Aspect)

When publishing data and information on the websites of authorities, a license for the content is usually not granted. This means that the content can be read, downloaded and used for personal use, but not reused.

This also applies implicitly, even if no information on this can be found in the imprint. Often, further restrictions are even provided, in the sense of “all rights reserved”.

Since one of the goals of open data is their reuse, an explicit granting of a free license according to the Open Definition<sup>1</sup> or a clear labeling of terms of use is required for their use. It is therefore necessary to decide (per dataset) which conditions are applicable, whereby the principle should apply: **As open as possible!**

Variants of Licenses	Note
<b>Public Domain</b> 	Intellectual creations that are not subject to intellectual property rights, in particular copyright, are considered public domain. <sup>2</sup> This also includes works without originality. The Public Domain Mark is not a license, but marks content that does not require a license. <a href="https://creativecommons.org/publicdomain/mark/1.0/deed.en">https://creativecommons.org/publicdomain/mark/1.0/deed.en</a>
<b>Zero (CC0)</b> 	Zero is the most open license. It means that the use of the data is not subject to any conditions. <a href="https://creativecommons.org/publicdomain/zero/1.0/deed.en">https://creativecommons.org/publicdomain/zero/1.0/deed.en</a>
<b>Attribution (CC BY)</b>  <b>Recommendation</b>	Attribution means that the use is subject to the only condition of naming the source. <a href="https://creativecommons.org/licenses/by/4.0/deed.en">https://creativecommons.org/licenses/by/4.0/deed.en</a>
<b>Other Licenses</b> 	There are numerous other licenses. Only use them if absolutely necessary. Further restrictions or “exotic” license variants make reuse or linking with differently licensed data more difficult.

Since there are no Austria-specific licenses, you can use the internationally known and recognized Creative Commons licenses.

CC0 is a recommended license, in Austria the CC-BY license has become the standard, where the indication of the source is mandatory.

Licensing open administrative data under a free license does not automatically mean that you have to change your information in the imprint of the entire website on which the data is offered. Deviating provisions may continue to apply to the content of the website. However, we recommend that you refer to the open data in the imprint.

<sup>1</sup> <https://opendefinition.org/od/2.1/de/> (Download 1. 3.2023)

<sup>2</sup> Since the assessment of what is public domain (what e.g. counts as a “free work” according to § 7 para. 1 UrhG or whether the data even reaches “originality” or “work character”) is not always easy, it

can happen in practice that content that is actually public domain is nevertheless marked as “Zero”. This may not be legally correct, but it is certainly in the interest of users, as they can rely on “Zero” to be able to use the data without restriction.



## Step 2: Machine-Readable Formats (Technical Aspect)

The second essential feature of open data is technically up-to-date formats. Data tables that have been published in PDF documents or on websites can be read using tools (via web “scraping”<sup>3</sup>), but this requires an unnecessarily high effort. You should therefore use formats that enable the easiest possible further processing.

Since in particular older applications may not be able to provide modern formats such as JSON, RDF or XML, CSV has become established as a minimum standard. Data in this format (Comma Separated Values) refers to tabular data that is stored in a text format and written one after the other separated by a delimiter (usually a comma or semicolon).

Spreadsheet programs such as Microsoft Excel or LibreOffice Calc can save such formats.

### TIP: Open licenses also for documents!

Although structured data is discussed here, it is advisable to apply the principles of free licenses to documents as well. For example, you can specify corresponding Creative Commons licenses in published documents. This guide, for example, is licensed under the Creative Commons CC BY 4.0 license and is published on [data.gv.at](https://data.gv.at)



The well-known 5-star model<sup>4</sup> serves to illustrate the importance of technical formats:

★	Publication of data on the internet under a free license (format irrelevant)
★★	Structured format (e.g. table instead of scanned image)
★★★	Open, non-proprietary format (e.g. CSV instead of Excel)
★★★★	Use of web standards (such as RDF) and URIs to make data addressable
★★★★★	Linking own data to other, addressable data (Linked Open Data)

### TIP: Publish what you have, but take export formats and APIs into consideration!

When procuring new IT systems or creating websites with prepared information, consider the question of reusability and machine readability right away: Request export formats such as CSV, JSON, RDF or XML.

Taking into account the provisions of the updated [Open-Data-Directive \(2019/1024\) Article 5](#) of the European Union, the provision of data via a dynamic API (Application Programming Interface; literally: application programming interface) is considered best practice. This involves interfaces in applications through which data can be accessed directly. APIs are the gold standard. Keep this option in mind, but don't postpone your publications until it is available, but publish immediately in the formats that are currently available to you.

<sup>3</sup> [https://de.wikipedia.org/wiki/Screen\\_Scraping](https://de.wikipedia.org/wiki/Screen_Scraping)  
(Download 1.3.2023)

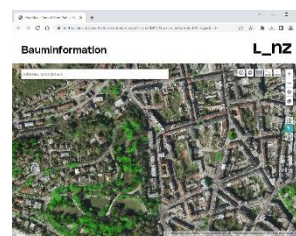
<sup>4</sup> <https://5stardata.info/de> (Download 1.3.2023)

### Step 3: Metadata (Organizational Aspect)

Metadata is “data about data”. Its aim is to describe the actually published dataset in more detail in order to make it more findable and understandable.

An example of metadata:

Title	Tree Register (Linz)
Description	All trees maintained by the City of Linz with information on genus, species, variety, height, crown diameter, trunk circumference, type and geo-coordinates.
Keywords	Trees, Register, Locations
Category	Environment & Climate
License	CC-BY
File Formats	JSON, CSV
Last Modified	February 24, 2023, 2:29 PM (UTC+01:00)
Publisher	City of Linz
Contact	Magistrat der Landeshauptstadt Linz, Stadtgrün und Straßenbetreuung (SGS)
Email	open.common@linz.at



The European standard for describing metadata is [DCAT-AP](#). Datasets whose metadata is described in this standard can also be forwarded to other data portals. The Austrian metadata standard is documented on [data.gv.at](#)<sup>5</sup> and is compatible with DCAT-AP.



<sup>5</sup> <https://go.gv.at/ogdmetade> (Download 1.3.2023)



# Which Data? – Data Catalog

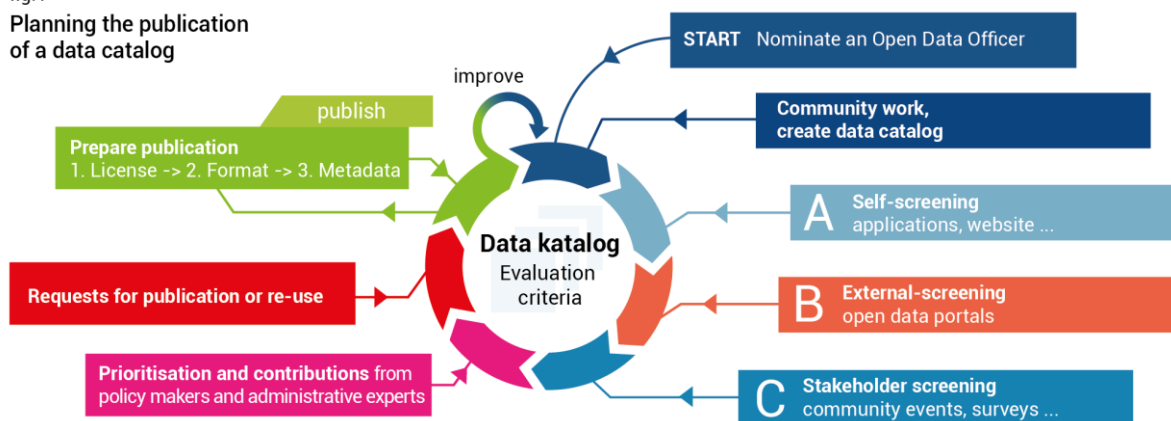
If you are planning and preparing the publication of data, internal responsibilities must first be clarified: nominate a person who coordinates the process as an [Open Data Officer](#). The focus of this coordination lies in internal and external [community work](#) and in creating and maintaining a data catalog. The internal community consists of contact persons in the various specialist departments, the external community (e.g. journalists, NGOs, developers) becomes relevant at the latest with the first publications. It is advisable to create an [internal data catalog](#) from the outset. This can record data that has already been published in another form, or data that is to be newly published. Three aspects must be considered:

- A. Self-screening:** What have we already published ourselves? Documents are often already published on the internet that do not yet meet the requirements of open administrative data. This offers the greatest potential to publish this data in an open format in the future or to offer new data in the open format with the next publication.
- B. External screening:** What have others already published? Here, a look at existing data portals such as [data.gv.at](#) or the European Data Portal [data.europa.eu](#) is particularly helpful. In Germany, a “reference data catalog”<sup>6</sup> is available, which describes datasets typically published by municipalities.
- C. Stakeholder screening:** What do the stakeholders want? Which data is of interest to different target and stakeholder groups? What data requests are received regularly? Are there already requests for open administrative data among them? Surveys and events in particular offer themselves as methods for getting to know the interests of stakeholders.

The resulting data catalog provides the basis for [prioritisations and contributions from policy makers and administrative experts](#). [Requests for publication or re-use](#) can also be answered with reference to open data, which can reduce administrative effort. Publication obligations for studies, expert opinions and surveys or other information subject to disclosure can also be incorporated into the data.

fig.1

Planning the publication  
of a data catalog



<sup>6</sup> [www.musterdatenkatalog.de](http://www.musterdatenkatalog.de) (Download 1.3.2023)

Once a sufficient decision-making basis has been created with the help of the data catalog to be able to make initial publications, the license and data formats are determined per dataset and the necessary metadata is collected. The dataset can then be published on a data portal and the process of revising the data catalog begins again. The evaluation criteria for internal data monitoring help with the assessment.

## Guidelines for Open Data

When creating a data catalog, questions arise such as: What is a dataset?" or "Which datasets are meant?" and "How detailed should we record?". The following guidelines are helpful here:

1. Data, not documents	Pure documents (e.g. files or other documents, PDF documents) are not data. But they can contain data. Therefore, it is about identifying and publishing the underlying data sources (e.g. Excel spreadsheet). In particular, in the case of studies, expert opinions and surveys, it will often be PDF documents that can be published as "Open Documents" on Data.gv.at. However, accompanying structured data should also be published here. In the case of surveys, this could be the (anonymized) raw results of an online survey.
2. Existing Data	The aim is not to collect new data or to compile data anew, but to publish the data that is already available in the organization.
3. Already published and unpublished data	In the case of data already published in another form, the license and the technical format must be taken into account in particular (see chapter "How to Proceed? – Steps to Open Data"). The assessment of whether previously unpublished data should be published is more complex. The criteria from internal data monitoring (see chapter "Criteria for Data Monitoring") are helpful here.
4. Data protection	Datasets that are inextricably linked to the names of natural persons can usually not be published. However, they could form the basis for an anonymized, pseudonymized or statistical dataset.

The Sunlight Foundation originally defined "ten OGD principles"<sup>7</sup>. The first three of these refer to the properties of the data to be published and are therefore also usable as guidelines:

1. Completeness	Datasets released should be as complete as possible, reflecting the entirety of what is recorded about a particular subject..
2. Primary sources	Datasets released should be primary source data. This includes the original informationt, details on how the data was collected and the original source documents recording the collection of the data. Public dissemination will allow users to verify that information was collected properly and recorded accurately..
3. Timeliness (currency)	Datasets released should be available to the public in a timely fashion. Whenever feasible, information collected should be released as quickly as it is gathered and collected. Priority should be given to data whose utility is time sensitive. Real-time information updates would maximize the utility the public can obtain from this information.

These principles are considered rough target values, not exclusion criteria for publication. If, for example, the complete dataset is not available in your own authority, but is available in another, you can certainly publish your own dataset, even if there is a more comprehensive one elsewhere.

<sup>7</sup> <https://www.data.gv.at/infos/open-data-prinzipien/> (Download 1.3.2023)

## Criteria for Data Monitoring

In particular when it comes to the question of whether data that has not yet been published can be published, the following criteria help with the assessment:

Criterion	Explanation
Non-disclosure/ legal restrictions	Is the data subject to confidentiality obligations or other legal restrictions or is it - data of critical infrastructure?
Personal reference	Is it personal data or can conclusions be drawn about persons from it?
Copyright	Does the administration have the sole right to use the data?
Value	How high is the estimated benefit for all target groups?
Effort	How high is the effort for publication?
Content-related data quality	How high is the estimated data quality? (Completeness, timeliness, accuracy, error rate ...)
Technical availability	Are the data formats and data sources available in open standards (OGD formats, 5-star model)?
Synergy	Is data/services already offered by the administration in another way? (There are bonus points here if publication is required anyway.)

These criteria can serve as a starting point for thought or can also be evaluated with points, e.g. from 0 to 5. The list of datasets with the highest point value is published as a priority.

### RECOMMENDED READING:

#### Open Government Data Screening in the Federal Government

Collection of potentially [open data capable datasets in the federal administration](#). A total of over 700 datasets were collected.



# What to Consider? – The Legal Framework

A constitutionally enshrined right to access information from the public sector in the form of a (internationally customary) Freedom of Information Act (IFG) is not yet in force in Austria, it is still in political consultation.

## Access to Information

However, there are still some legal provisions that regulate access to information. The basis is Article 20 of the [Federal Constitution](#), which contains three essential paragraphs:<sup>8</sup>

- Art 20 para. 3 B-VG defines [official secrecy](#) in certain cases, which is only permissible as an argument for refusing to provide information in a narrow framework..
- Art 20 para. 4 B-VG therefore immediately follows the [obligation to provide information](#), which is also reflected in the information obligation laws of the federal government and the states.
- Art 20 para. 5 B-VG was newly introduced on 1 January 2023 and concerns the [publication obligation](#) for [studies](#), [expert opinions](#) and [surveys](#) commissioned externally by administrative bodies, including their costs.

In addition, there are publication obligations in various other legal areas, e.g. for traffic data ([IVS-RL](#)), geodata ([INSPIRE-RL](#)), environmental data ([Environmental Information Directive](#)), procurement data ([BVerG 2018](#)) or financial data ([Stability Pact 2012](#), [VRV 2015](#)).



<sup>8</sup> Federal Constitution <https://www.ris.bka.gv.at/eli/bgbl/1930/1/A20/NOR40245775> (Download 11.3.2023)

## Framework Conditions for Further Use

The [Open Data Directive 2019/1024](#)<sup>9</sup>, which is a recast of the PSI Directives that already existed in 2003 and 2013, does not fundamentally regulate access to information, but rather the framework conditions for further use, such as fees, formats, application procedures and licenses. They have been implemented in Austria as the [Federal Information Reuse Act \(IWG\)](#) and [Information Reuse Acts](#) of the states.<sup>10</sup>

A special case is so-called “[high-quality datasets](#)”<sup>11</sup>. The European Commission obliges the member states to make these available for reuse by June 2024, and to do so via APIs (application programming interfaces) or mass download.

Categories of such data are:

Geodata	Administrative units, geographical names, addresses, buildings, cadastral plots, reference plots
Earth observation	Water network, protected areas, elevation, geology, land cover, orthophotography, management areas, biogeographical regions, energy sources, environmental monitoring facilities, habitats and biotopes, land use, mineral resources, areas with natural hazards, oceanographic-geographic parameters, production and industrial facilities, marine regions, soil, species distribution
Environment	Air, climate, emissions, nature conservation and biodiversity, noise, waste, water
Meteorology	Observation measurements from weather stations, climate data: validated observations, weather warnings, radar data, model data from numerical weather forecasts
Statistics	Industrial production, industrial producer price index, sales volumes by activity, statistics of international goods traffic, tourism flows, population, fertility, mortality, national accounts, government expenditure and revenue, consolidated gross government debt, poverty rate, inequality rate, employment, unemployment, potential labor force
Companies and ownership of companies	Basic company information, company documents and financial statements
Mobility	Transport networks, inland navigation



<sup>9</sup> Directive (EU) 2019/1024 <http://data.europa.eu/eli/dir/2019/1024/oj> (Download 11.3.2023)

<sup>10</sup> Information Reuse Act 2022 of the Federal Government <https://www.ris.bka.gv.at/eli/bgbli/ii/2022/116/P0/NOR40246446> or overview of state legislation n <https://www.bmaw.gv.at/Themen/Europa/Wettbewerbspolitik/Open-Data.html> (Download 11.3.2023)

<sup>11</sup> Implementing Regulation (EU) 2023/138 of the Commission of 21 December 2022 [http://data.europa.eu/eli/reg\\_impl/2023/138/oj](http://data.europa.eu/eli/reg_impl/2023/138/oj) (Download 11.3.2023)

## Open by Default

The increasingly complex regulations highlight two things:

Firstly, legal norms – not least driven by EU regulations and current case law – are moving towards greater openness, more comprehensive publication obligations and stricter regulations on further use.

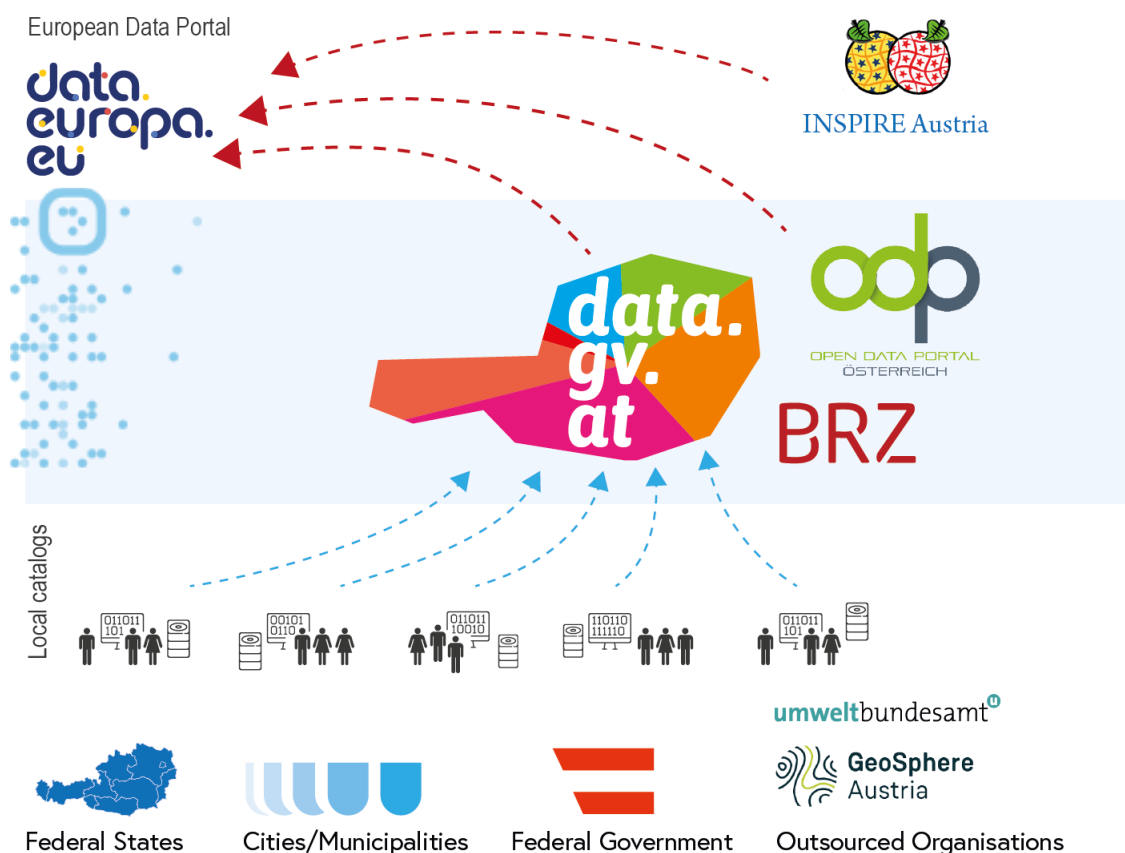
Secondly, proactive, voluntary publication as open data is therefore often the option with the least additional effort. Therefore, the principle of “[Open by Default](#)” should apply: public bodies use proactive publication as open data wherever possible, this saves internal effort and the more complex legal considerations are then left to more difficult cases in terms of content.

Question	Regulatory area	Explanation
Official secrecy or freedom of information	Access	Do the following, narrowly construed confidentiality criteria apply? <ul style="list-style-type: none"> <li>the interest in maintaining public peace, order and security;</li> <li>the interest in comprehensive national defense;</li> <li>the interest in foreign relations;</li> <li>the economic interest of a public law corporation;</li> <li>the preparation of a decision;</li> <li>the overriding interest of the parties.</li> </ul> If no, information must be provided or proactive publication can take place.
Data protection	Access and processing	Does the data contain personal data and do the data subjects have a legitimate interest in keeping it confidential? If yes, can the personal data be removed or anonymized?
Study, expert opinion or survey	Access	Were studies, expert opinions or surveys commissioned from external parties after 1 January 2023? If yes, <ul style="list-style-type: none"> <li>publish the PDF document as an “Open Document” on Data.gv.at (and carry out redactions if necessary).</li> <li>information on the costs can be stored in the “Description” metadata field.</li> <li>publish the underlying structured data (e.g.: survey results) as well. (If necessary, remove personal references, e.g. IP addresses or email addresses in survey raw data.)</li> </ul>
IWG reuse application	Reuse	Should the reuse application be granted? If yes: Instead of individual responses, it is more efficient (if possible) to publish the requested data on Data.gv.at and to refer to it in the response. This also saves future requests, as the data can then be reused anyway.
High-quality datasets	Access and reuse	Implementation required by June 9th 2024. Requirements: <ul style="list-style-type: none"> <li>free of charge</li> <li>machine-readable,</li> <li>available via application programming interfaces (API)</li> <li>available as mass download if necessary.</li> </ul>

## Where to Publish? – Data Portals

The publication of a dataset can take place on your own website. The most important criterion is that the dataset is permanently accessible via a URL. The dataset from our example (Tree Register Linz) is accessible under its own URL<sup>12</sup>.

Data portals are not absolutely necessary for publication. You can also offer datasets for download on your own website, specify the corresponding metadata there and define a license. Portals such as [data.gv.at](https://data.gv.at) or the European Data Portal are so-called [metadata directories](#). Their purpose is to record datasets and make them more findable. Such metadata portals usually also offer the possibility of uploading datasets to make them accessible there. However, it is generally sufficient to describe the metadata there and to specify another source of the datasets (e.g. the web server of your authority).



The Austrian data portal [data.gv.at](https://data.gv.at) is available to all public institutions for free use. Once collected metadata is forwarded from one portal to the next (via so-called “harvesting”) and passed on to the European data portal “[data.europa.eu](https://data.europa.eu)”.

<sup>12</sup> <https://data.linz.gv.at/katalog/umwelt/baumkataster/> (Download 1.3.2023)



## How to Start? – Phases of Implementation

The following overview helps you to find your way through the phases that you will probably go through in the course of publishing open data:

	No Open Data	Phase I: Beginner	Phase II: Advanced	Phase III: Professional
Data Focus	none	already published	unpublished	open by default
Process	Website	ad hoc	established	evaluated
Departments involved	as needed	willing	important	all
Influence on the organization	–	sporadic	regularly	strategically
Community involvement	reactive (response to requests)	proactive: data requests, surveys	shaping: hackathons, events	continuously
Usage figures	–	record	analyze and publish	adapt data offering
Data formats	PDF/Excel	CSV	XML/JSON	RDF
Geo-reference	–	CSV, coordinates	GeoJSON, raster data	WMS, vector data, OGC interface
Publication	manually	manually	export (from applications)	API
Metadata recording	none	manually (via form in the data portal)	import (into the data portal)	API (harvesting)
Data evaluation (according to criteria)	none	occasional (personal, email)	structured (Excel)	continuously (OGD cockpit)
Reference data catalog	–	top topics (What do almost all have? E.g. public transport, infrastructure...)	potentials (What do others have that we are still missing?)	complete
View of others	irrelevant	regional	national	international
Expectations (external)	information (no reuse)	attention, prototypes	apps, applications	startups, business models
Expectations (internal)	–	awareness	benefit	change
Goal level	–	publish data	optimize publication processes	data governance
Open-Government-Implementation Model <sup>13</sup>	–	1 Transparency	2 Participation and 3 Collaboration	4 Sustainability and Public Value
Open Data Maturity Level <sup>14</sup>	–	1 Initial -> 2 Repeatable	3 Defined -> 4 Managed	4 Managed -> 5 Optimizing
5-star model <sup>15</sup>	* to **	***	****	*****

Those who do not publish open data provide information on the authority's website on a case-by-case basis, do not care about licenses or formats, and thus make reuse more difficult. The potential of open data therefore remains unused. On request, you would refer to PDF documents on your website in this case or occasionally send an Excel file by email.

In the **beginner phase**, you try to publish open data once. Your focus is therefore on data that has already been published in another form

You look for willing departments and start to proactively take into account the wishes of stakeholders. The first publications can be done manually (both data and metadata are uploaded individually). You only carry out data evaluations, if at all, in occasional discussions with departments.

<sup>13</sup> Phases according to Krabina and Lutz 2016

<sup>14</sup> <https://theodi.org/article/open-data-maturity-model-2> (Download 11.3.2023)

<sup>15</sup> <https://5stardata.info/de> (Download 11.3.2023)

You orient yourself in particular with comparable authorities and network with them. You expect to generate initial attention for the topic (internally and externally), first prototypical data uses will help you with this.

As an **advanced organization**, you convince more and more important departments in your organization, who may have been sceptical before. You need to optimize your publication processes, as the manual publication approach will soon no longer be sufficient. You manage to have departments regularly generate new datasets (e.g. by exporting from applications) and enter metadata. In order to maintain an overview and to be able to plan publication phases, you now carry out data evaluations in a more structured way. In other data catalogs, you increasingly look at the potentials (i.e. at datasets that others have already published, but you have not yet). You therefore orient yourself more nationally (at other authorities).

First apps and applications are emerging that are driving your initiative forward.

As **open data professional**, an open by default approach has become established. IT systems are equipped with open APIs for data and metadata. You orient and network internationally, you are in ongoing contact with the emerging startup scene. Internally, you pay attention to data governance, in which all departments are involved. Many internal processes have therefore changed and improved.

The phases described here are only described as examples and are probably hardly ever encountered in practice in this form. The table is intended in particular to provide orientation and an overview of which aspects of your own OGD initiative you could still improve.

#### RECOMMENDED READING:

##### Open Government Implementation Model

The [Open Government Implementation Model](#) of the KDZ – Center for Administrative Research contains numerous further recommendations for the implementation of open data. In particular, the further phases of Open Government are also described in detail.



100 01100001 00101110 01100111 01110110 00101110 01100001 01110100

00101110 01100111 01110110 00101110 01100001 01110100

00101110 01100001 01110100 01100001 00101110 01100111 01110110 00101110 01100001

01100100 01100001 01110100 01100001 00101110 01100111 01110110 00101110 01100001 01110100



---

# Why Open Data? – Open Government

In conceptual terms, open data is a key element of the first phase of Open Government: transparency. The further phases are participation and collaboration. These phases are interconnected and build on each other. Open Government seamlessly connects to concepts of public management and governance. Open Government can therefore be understood as a comprehensive redesign of political and administrative action in the sense of modern public management and public governance.

When answering the question “Why open data?”, different aspects play a role. It is advisable to differentiate and distinguish these aspects in the discussion <sup>16</sup>.

## Technical Arguments

- Data should not be “locked up” in documents. Machine-readable data facilitates both external and internal use. You could also consider this better accessibility of data..
- Modern and up-to-date formats such as JSON and RDF facilitate reuse.
- Standardized metadata allows descriptions of datasets to be exchanged between data portals (“harvesting”). Metadata recorded on data.gv.at is automatically forwarded to the European Data Portal.
- APIs enable the timely use of data (real-time data).

## Economic Arguments

- The Open Data & PSI Directive was created because it was recognized in Europe as economically disadvantageous if data from the public sector is not accessible. The Open Data study by Danube University Krems dealt with the economic added value for Austria.<sup>17</sup>
- Open data is intended in particular to stimulate the local economy, as numerous apps and applications are emerging that do not have to be financed by the organizations. They are intended to contribute to the emergence of new ideas and business models.
- Often, there are “revenue illusions” in authorities, e.g. when the focus is only on the revenue from data sales, without comparing the total costs that would arise for marketing, sales, billing, etc. of the data.<sup>18</sup>

## Legal Arguments

- The central legal argument is the legal certainty that is created by granting open licenses.
- Numerous legal frameworks (EU directives and national laws) have strengthened the argument for proactive publication of open administrative data in recent years. (See chapter “What to Consider? – The Legal Framework”)
- The General Data Protection Regulation (GDPR) also plays a role, which calls for the special protection of personal data. Personal data is not the subject of open data by definition, but this is a too narrow view, as data in which personal references occur or can be established can also be useful.

## Political Arguments

- Political arguments include in particular the fact that data transparency is an important basis for participation. But not only “the involvement of citizens” needs open data, evidence-based policymaking is also increasingly back on the agenda in the age of Big Data.
- A political question in particular is also which data should be published that was not previously accessible at all.

---

<sup>16</sup> A more detailed discussion of the different aspects can be found in Krabina 2019.

<sup>17</sup> See Viale Pereira et al. 2017

<sup>18</sup> See Jörg 2014

## Societal Arguments

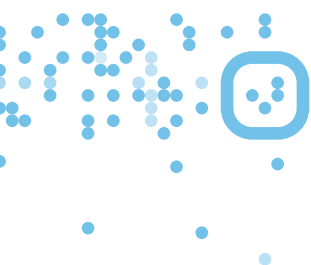
- The creation of the data was already paid for with tax money, and therefore data should be made available to the public without restrictions.
- From the perspective of classic public services (public services that constitute a basic supply), one can also argue that several new tasks become important in the course of digitization as digital public services. In addition to providing a digital basic infrastructure (broadband internet, WLAN hotspots, IT equipment) and enabling new services (public space servers, labs), this also includes open data.
- Current developments such as Data Governance, Data Science and Data Journalism, Big Data and Data Analytics, Smart Cities and Smart Government or Artificial Intelligence benefit from an increasing number of freely available data from the public sector.
- The availability of underlying data increases trust in the information of the public sector.<sup>19</sup>

## Organizational Arguments

- Released data is also used within the administration.
- The intensive discourse on data with users is seen as enriching by many public administration employees.
- The fact that a dataset is to be published often leads to an increase in data quality, as metadata descriptions have to be created (to explain the structure of a CSV file, for example) or errors in published data are reported.
- Authorities that publish their data on open data portals receive significantly fewer requests and applications for access to public files than authorities that do not publish their data. This means that authorities that regularly publish data can save time.<sup>20</sup>

## Administrative-Economic Arguments

- Fundamental questions, such as the collection, further processing and publication of data, should be completely reassessed in light of today's available technologies. Here, the distribution of data competencies between municipalities, states and the federal government would also have to be questioned.



### RECOMMENDED READING:

#### Economic and Societal Potential of Open Data



The [Study](#) the University of Continuing Education Krems examines the social, political and economic impacts of open data and provides recommendations for politics and administration.

<sup>19</sup> See Schmidhuber, Willems and Krabina, 2023

<sup>20</sup> See Bruns, Dittwald and Meiners 2019

# What to Expect? – Frequently Asked Questions / FAQ

## Does Open Data require a legal basis?

Generally no. Just as an administration decides for itself which information it publishes on its website, it can also decide for itself to publish data. No new legal framework is required for this. In addition, it should be noted that some areas of law, in particular EU directives (see above “Legal Arguments”), are clearly developing towards open data. The revised PSI Directive is now called, for example, Open Data and PSI Directive (EU) 2019/1024. You can therefore start with open administrative data at any time, even if there is no Freedom of Information Act yet. (See chapter “How to Proceed? – Steps to Open Data”.)

## Does Open Data contradict data protection?

No. Personal data is not the subject of open data by definition. This is usually data that does not have any personal references (see above, “Criteria for Data Monitoring”) and whose processing and publication is not subject to the General Data Protection Regulation (GDPR). However, this is a too narrow view, as data in which personal references occur or can be established can also be useful. Therefore, methods of anonymization or synthesis must be applied to data with personal references before it can be published.<sup>21</sup> (See chapter “Which Data? – Data Catalog”.)

## If we make the data completely open, will we lose control?

Yes, but in many areas we have already lost control anyway! Much of the data from the public sector is collected via commercial providers such as Google (Google Maps) or free community projects such as Wikidata or OpenStreetMap, often because the public sector was too slow to offer the data itself. If data is freely available to everyone, it can be used by both commercial and non-commercial services, with official data supplementing existing offerings. (See chapter “Why Open Data? – Open Government”.)

## Shouldn't we sell data instead?

No. While it is conceivable to offer a basic dataset as an open and free dataset and a more comprehensive one for a fee, there are several arguments against authorities selling data. There are revenue illusions (the effort involved in selling data is underestimated) and customers switch to alternative products. (See chapter “Why Open Data? – Open Government”.) In addition, according to the Open Data and PSI Directive (EU) 2019/1024, the charging of fees is only permissible in a very narrow framework (maximum marginal costs).



<sup>21</sup> Siehe dazu Drechsler und Jentzsch 2018.

## But we have already published it as a PDF, why the double effort?

Double effort is only an artifact of the beginning phase. Once processes have been established where data from specialist applications is automatically published, the effort will decrease again. And consider how much effort the various departments put into producing printed publications, such as a statistical yearbook, where the benefit is “only” in informing the interested public – without the added benefit of reusability.

Consider offering open data as a form of accessibility: data is “locked up” on websites and in PDF documents and is made accessible through opening. (See chapter “How to Proceed? – Steps to Open Data”.)

## We don't have any (relevant) data!

Yes, you do. A look at data catalogs from others helps to get a quick overview of data that could also be available in your authority. There are different ways to use it. Data publications do not always lead to groundbreaking new applications. But informing the public is always a relevant and important task of authorities. Often there is also a lack of demand orientation: Which data is actually of interest and in demand? Hackathons and other events to involve various stakeholders help with this (See chapter “Which Data? – Data Catalog”.)

## Just putting out data doesn't do anything. How do we get apps?

It is advisable to maintain continuous exchange with all stakeholders. The sooner you know which data they consider relevant, the faster your data will be used. (See chapter “How to Start? – Phases of Implementation”.)

Data.gv.at lists more than 700 apps and applications



## Do we have to operate our own data portal?

No. In Austria, data.gv.at is a metadata portal available to all institutions in the public sector. (See chapter “Where to Publish? – Data Portal”.)

## We don't have the rights to this data!

Try to clarify any missing rights. Often, such questions are forgotten in advance. Therefore, when you next tender a project (e.g. study, publication or website) or a software solution, remember that open data should be made available. This saves you work and also serves to clarify the legal aspect. (See chapter “Which Data? – Data Catalog”.)

## What is the difference between “FAIR” and “open”?

The “FAIR data principles” were developed with the aim of improving the use of research data, according to which (research) data should be “Findable”, “Accessible”, “Interoperable” and “Reusable”.<sup>22</sup> The FAIR principles can also be fulfilled if actual access to the data is subject to a fee or registration. FAIR therefore does not yet mean open. Conversely, open data generally fulfills the FAIR principles, but in detail, differences can arise, in particular with regard to “interoperable”. How interoperable open data is depends in particular on the technical format (see 5-star model).

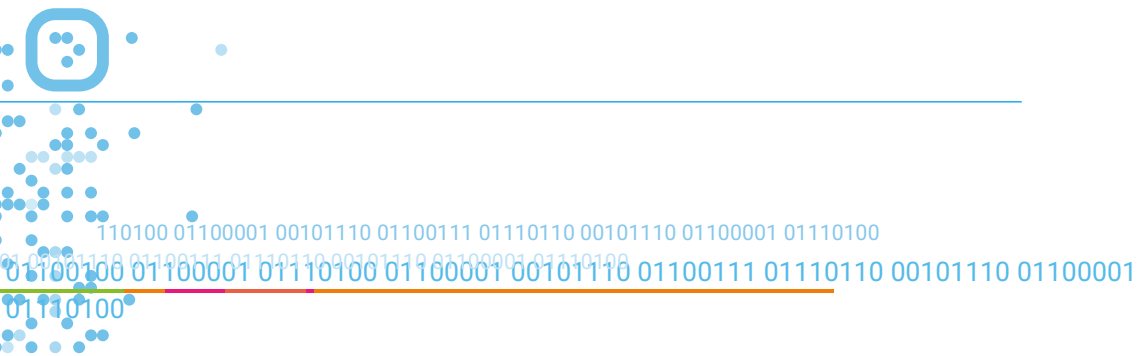
## Is open data harmonized?

No. Open data is offered in the structures in which it is available in the authorities. These structures may have been standardized at either national or international level, in which case there is a substantive harmonization of data. Open data can therefore be harmonized, but this is often not the case.

<sup>22</sup> Further information see <https://forschungsdaten.info/fdm-im-deutschsprachigen-raum/oesterreich/>

### How can the use of sensitive data be enabled?

It may happen that data that contains personal references or cannot be made openly available for other reasons should nevertheless be made available for further use. The European Union wants to create framework conditions for this with the Data Governance Act. Also, “data spaces” are to be created to make non-public data (e.g. business-relevant data) securely usable under clear contractual and trustworthy conditions. The European Data Space for Health Data (European Health Data Space EHDS) is an example of this.

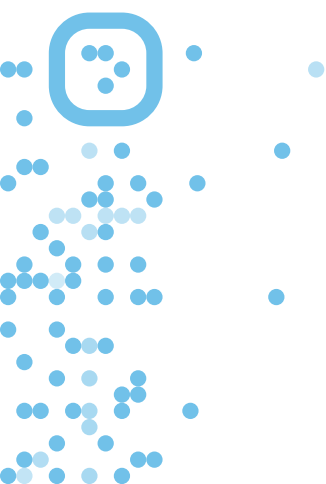




---

## Which Terms Should You Know? – Glossary

API	Application Programming Interface. Interface for the dynamic provision of data.
CC-BY	Creative Commons License. BY means Attribution.
CSV	Comma-Separated-Value. Tabular data separated by delimiters.
DCAT-AP	European standard for describing metadata.
JSON	JavaScript Object Notation. Data format in a particularly machine-readable text form.
LOD	Linked Open Data.
Metadata	Data that describes datasets.
OGD	Open Government Data. Open administrative data.
RDF	Resource Description Framework. Data format for describing resources on the internet.
URL	Uniform Resource Locator. Locates a resource, for example a website.



# Where Can You Read More? – Further Information

## Guides and Studies

- Data.gv.at - Guides
  - „Open Data Analysis – Making Better Decisions“ (2020)
  - „Open Data Governance – On the Way to a Data-Driven Organization“ (2020)
- Cooperation OGD Austria: Publications on framework conditions or Austrian metadata standards:  
<https://www.data.gv.at/infos/cooperation-ogd-oesterreich/>
- Data.europa academy: English training materials on open data
- Bertelsmann Stiftung: [A Guide to Open Data](#) (2020). Version of this guide for Germany. Contains further information on guides and studies from Germany as well as an overview of software for data portals.
- BKA: [OGD Screening in the Federal Government](#) (2017). Results of a project conducted in 2017 to collect potentially OGD-capable datasets in the federal government.
- Danube University Krems: [Economic and Societal Potential of Open Data](#) (2017)
- KDZ – Center for Administrative Research: [Open Government Implementation Model](#) (2016)

## Internet Links

- Austrian Data Portals: [data.gv.at](https://data.gv.at) (offene government data), [opendataportal.at](https://opendataportal.at) (open data from business and civil society), [mobilitydata.gv.at](https://mobilitydata.gv.at) and [inspire.gv.at](https://inspire.gv.at) (mobility and geodata, respectively), [offenerhaushalt.at](https://offenerhaushalt.at) (financial data) und [forschungsdaten.at](https://forschungsdaten.at) (research data).
- European Data Portal: <https://data.europa.eu/en>
- 5-Star Model: <https://5stardata.info/>
- Data Intelligence Offensive (DIO): [Dataintelligence.at](https://Dataintelligence.at)



---

## Bibliography

- Drechsler, Jörg, und Nicola Jentzsch (2018).  
"Synthetic Data. Innovation Potential and Societal Challenges". Ed. Stiftung Neue Verantwortung. Berlin. [https://www.stiftung-nv.de/sites/default/files/synthetische\\_daten.pdf](https://www.stiftung-nv.de/sites/default/files/synthetische_daten.pdf) (Download 1.3.2023).
- Groß, Marc and Bernhard Krabina (2019).  
"Data-Centric Administrative Management". Sustainable Impact. Impulses for the Public Sector. Ed. Peter Biwald, Karoline Mitterer, Thomas Prorok und Klaus Wirth. Reihe Öffentliches Management und Finanzwirtschaft, Band 22. Wien: NWV Verlag.
- Jörg, Wolfgang (2014).  
"ViennaGIS® Gives Away Its Geodata – Can We Afford It?". Surveying & Geoinformation 3. 138–145. <https://www.ovg.at/de/bibliothek/vgi-die-zeitschrift/files/pdf/5180> (Download: 1.3.2023).
- Krabina, Bernhard (2019).  
"Open Government and Open Data as Modernization Concepts: Opportunities and Challenges of Open Administrations". E-Government Handbook. Technology-Induced Administrative Development. Ed. Jürgen Stemberger, Wolfgang Eixelsberger, Alessia Neuronì und Andreas Spichinger. Wiesbaden. (DOI: [10.1007/978-3-658-21596-5](https://doi.org/10.1007/978-3-658-21596-5)).
- Krabina, Bernhard, und Brigitte Lutz (2016).  
„Open Government Implementaton Model”.  
<https://www.kdz.eu/de/open-government-vorgehensmodell>  
(Download: 1.3.2023).
- Miernicki, Georg (2022).  
"The Publication Obligation of Information of Administrative Bodies". ÖJZ 2022/158.
- Schmidhuber, Lisa, Jorgen Willems and Bernhard Krabina (2023).  
"Trust in public performance information: The effect of data accessibility and data source." Public Administration Review 01/2023. DOI: [10.1111/puar.13603](https://doi.org/10.1111/puar.13603).
- Viale Pereira, Gabriela, Shefali Virkar, Valerie Albrecht and Peter Parycek:  
"Economic and Societal Potential of Open Data. Report on the Impacts of Open Data V2". Vienna, 2017. <https://www.data.gv.at/katalog/dataset/ff51457a-2c9f-4553-b5f0-ea10a0c71119> (Download: 1.3.2023).



## Imprint

Cooperation OGD Austria

### Author

Bernhard Krabina

COOPERATION OGD  ÖSTERREICH

Open Commons 



The text of this publication is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) License.

The full license text can be found at:

<https://creativecommons.org/licenses/by/4.0/legalcode.de>



All *photos* and *logos* are excluded from this, they are not subject to the above CC license.

This is the English version of the document available at <https://data.gv.at/leitfaden>. It was translated using AI and checked by the author. No other updates were made.

The English version was co-funded by the INTERREG project "DoorCE – Data Opener Central Europe" <https://www.interreg-central.eu/projects/doorce/>

**Interreg**  
CENTRAL EUROPE



Co-funded by  
the European Union

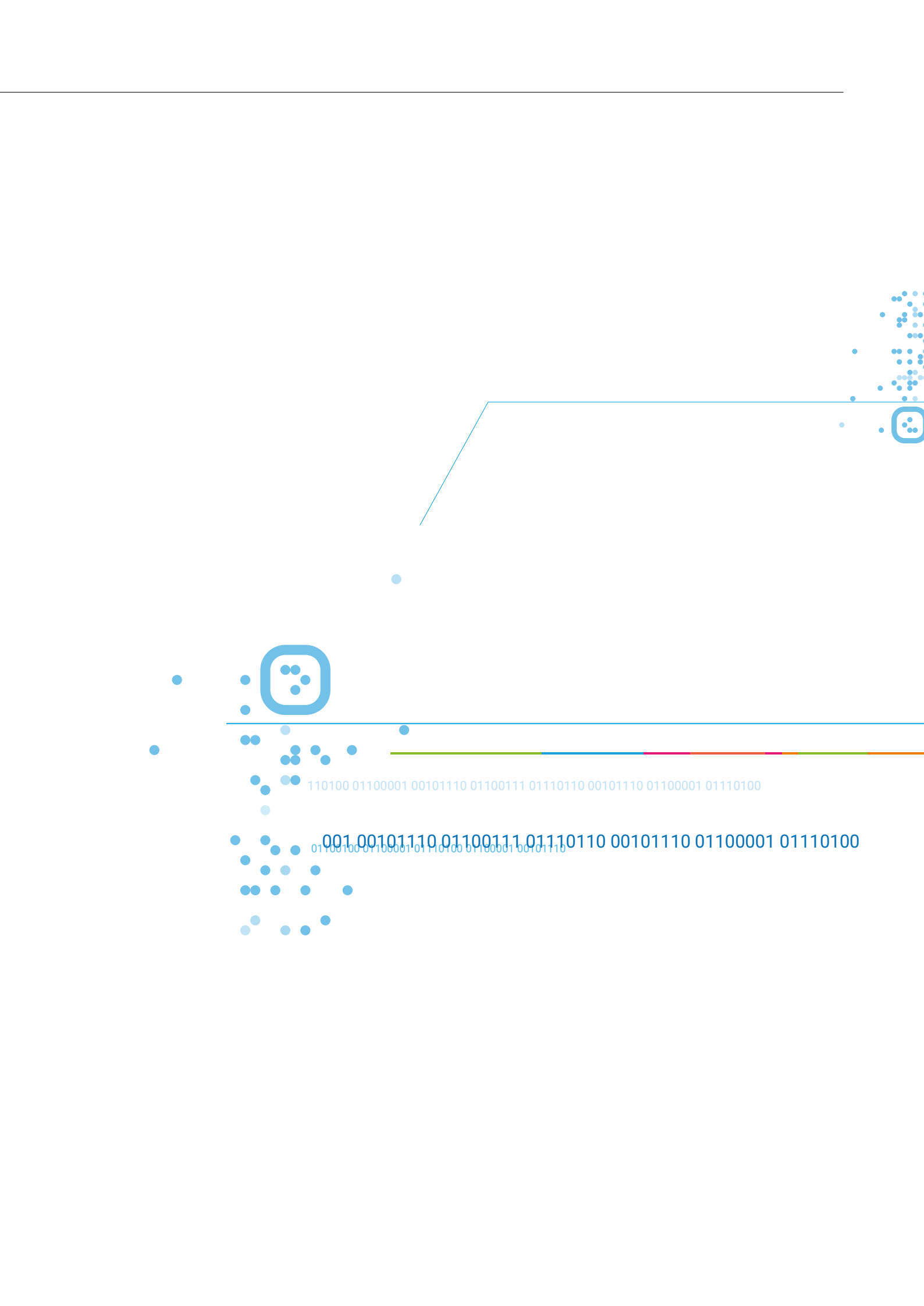
DoorCE

### RECOMMENDED READING:

#### Guide to Open Data of the Bertelsmann Foundation

The "Guide to Open Data ("Ein Leitfaden für offene Daten") of the Bertelsmann Foundation" is the version of this guide that is tailored to the situation in Germany.

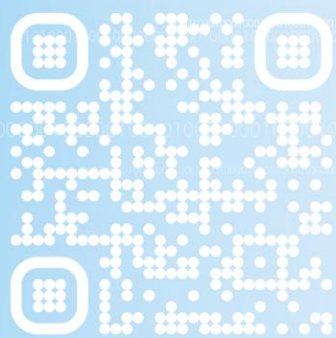




110100 01100001 00101110 01100111 01110110 00101110 01100001 01110100

001 00101110 01100111 01110110 00101110 01100001 01110100

01100100 01100001 01110100 01100001 00101110



## Contact

Bernhard Krabina  
bernhard.krabina@ikt.linz.at



Co-funded by  
the European Union

DoorCE

