

Framework for Open Government Data platforms		White Paper
		Open Government Data – 1.1.0
		Result of PG
Short description	<p>The disclosure of data management is discussed as a means to increase participation in a common value creation process of politics, administration, citizens and economy. The accompanying transparency of data and information increases the trust of all stakeholders in administrative processes, leads to new business models and strengthens in the medium to long term the existing democratic institutions.</p> <p>This document provides basic information available about Open Government Data and defines legal, technical and organizational requirements for Open Government Data platforms.</p>	
Authors:	Gregor Eibl (Federal Chancellery) Johann Höchtl (Danube University Krems) Brigitte Lutz (Vienna) Peter Parycek (Danube University Krems) Stefan Pawel (Linz) Harald Pirker (Federal Chancellery)	Project Team / Working Group Project Group Open Government Data Austria
Translation:	Gregor Eibl (Federal Chancellery)	

Table of Contents

1.	Basic information about Open Government Data	3
2.	Open Government Data Principles.....	5
3.	Uniform name for administrative management.....	8
4.	Technical requirements - Open Government Data Formats & Metadata ...	8
5.	Organizational Requirement - OGD monitoring.....	9
6.	Legal Requirement – Licensing and Terms of Use.....	11

Open Government Data in the public administration

1. Basic information about Open Government Data

Open Government (opening of state and government)

Open Government is used as a comprehensive term for a range of concepts and visions that deal with certain aspects of opening of state and government. These include considerations about Transparency 2.0, Participation 2.0 and Collaboration 2.0, the open innovation approach, opening of the society, open society, reflections on free data and open standards, open interfaces, open source software and open communication systems."¹

The principles of Open Government are:

- **Transparency:** strengthens the sense of duty and provides the citizens information about what their government and their administration do. The free availability of information is an essential basis for transparency.
- **Participation:** strengthening the effectiveness of government and administration and to improve the quality of their decisions by the widely dispersed knowledge of the society involved in the decision making.
- **Collaboration:** new tools, methods and systems to accelerate the cooperation of all administrative levels as well as those with the private sector.

An implementation of an Open Government strategy in administrative and political processes should therefore be made in the light of the transformation of values and decision-making cultures. For this purpose transparent processes are necessary that make Open Government Data become a key pillar of all Open Government strategies.

¹ von Lucke, Open Government, 2010, P. 3

Open Government Data (OGD)

Open Government Data are those non-personal and not infrastructure-critical data resources held by public bodies that are made accessible in the interest of the general public without any restriction for free usage, for dissemination and for re-use.

Open Government Data is awarded the potential for sustainable social, cultural, scientific and economic progress. Through the utilization of non-personal information in the public sector new products and services will be promoted and economic growth will be supported in Austria. Moreover, Open Government Data is seen as an adequate tool to increase the transparency of administrative action, to allow better cooperation between politics, government, business, research and citizens and to strengthen democracy.

In selecting the data to be published, care should be taken that those data are selected which are interesting and useful for users.

Potentials

The key opportunity for the state is the use of the innovative power of society and economy. Transparent data and accountable decision-making processes are key for voluntary participation in joint value creation processes.

Applications based on Open Data can lead to better and more efficient services. Positive societal effects are expected due to indirect profitability through new business models and thus higher tax revenue. The precise utility and value of Open Government Data for policy and administration can be quantified only vaguely.

For administration other economic opportunities arise besides the "outsourcing" of application development by tax revenues drawn from sold Open Data-based products.

With the establishment of an Open Data strategy the interoperability of data management is improved and enables the linking and analysis of internal databases. Internal records are visible to both internal and external stakeholders and thus promote cooperation. The opportunity lies in a variety of analyses that can lead to new insights.

The integration of society and the economy can improve data quality using crowdsourcing methods.

In addition to the economic advantages democracy is strengthened by increased transparency, which potentially leads to a strengthening of citizens' trust in the state and its institutions.

Risks and obstacles

Besides the above-mentioned positive effects negative ones are also to be expected by higher transparency. One of the biggest risks of Open Data are the aforementioned errors of interpretation of data; statistical data, for example due to lack of expertise, lack of background knowledge or by the mixture of data sets of different quality can be misinterpreted. Quality assurance is therefore a critical factor that could both top-down carried out by the state and its institutions, as well as bottom-up, by means of examination and counter analysis by the society, especially through science and involvement of universities and colleges.

Open data lead to more transparency which can, however, have negative political and social, individual effects. For example, an easier access to environmental data through map-based representation, such as soil contamination of land, would lead to direct adaptations in the real estate market.

The effect of transparency on audited data is undisputed in this regard; the critical question is how politics, administration and society deal with the results. Even the deliberate manipulation by data corruption in statistical analysis or falsified representation in applications, creates new challenges for the administration offering data.

By publishing Open Government Data additional initial and maintenance costs (quality, availability...) need to be considered when planning budgets.

Even if Open Government Data - records are never personal datasets, there exists a certain risk that an assignment to individuals cannot be excluded by merging multiple data sources. This possibility has to be considered and assessed in advance.

2. Open Government Data Principles

Eight Open Government Data principles were already formulated in late 2007 by Open Government advocates in the United States, enhanced by the Sunlight Foundation to ten principles².

² OpenGovData.org, 2007. 8 Principles of Open Government Data: <http://www.opengovdata.org/home/8principles>
Translation of the Ten Principles for Opening up Government Information from the Open Data Network

Whenever a public body publishes data as Open Data, the following principles should be adhered to as far as possible. If the compliance of individual principles is not possible, this should be justified internally and externally documented.

1. **Completeness:** the published datasets are as complete as possible; they form the entire extent of what is documented on a specific topic. Metadata describing the raw data are also included along with formulas and explanations to calculate the data. This will allow users to understand the adjustments of available information and to examine each data element with the greatest possible detail. Before publication data protection restrictions, security restrictions or access restrictions have to be checked. Personal data are basically exempt from the publication.
2. **Primary source:** The data is collected at its origin and is published with the finest possible granularity, not in aggregated or otherwise modified formats.
3. **Current provision:** the published records are available to the public within an adequate period as topically as possible. They are published as soon as they are collected and compiled. Data available in real time are retrievable through a programming interface (API).
4. **Easy access:** the published datasets are very easily accessible without barriers. Physical barriers (e.g. the need to personally go to a certain office or the requirement to meet certain processes) must be avoided, such as technical barriers (e.g. access to data only after completed entry forms or via systems requiring browser-based technologies such as Flash, JavaScript, Java applets or cookies).
5. **Machine-readability:** data are stored in established file formats that are easily machine-readable so that an automated, structured processing is possible. The use of different file formats is recommended. If for any reasons the use of non-machine-readable formats is required, the data should also be available in machine-friendly formats. Files should be accompanied by a documentation which refers to the format and on how it can be used in respect to the data.
6. **Discrimination:** Any person may at any time access the data without identifying him-/herself or having to give a justification for her actions.
7. **Use of open standards:** The formats in which the data are published are as open as possible, over which no entity has exclusive control (see "OGD Formats" in the White Paper OGD metadata). Here, the administration orientates on standards that were developed by bodies such as the World

Wide Web Consortium (W3C), and at conventions of the Austrian BLSG or recommendations of the SAGA³ in Germany.

8. Licensing: The datasets are published under the license Creative Commons Attribution 3.0 Austria (AT CC BY 3.0) <http://creativecommons.org/licenses/by/3.0/at/deed.de>. The administrative unit must clarify any copyright, patent and trademark questions in advance.
9. Documentation (durability): Information published by the administration is fully documented with metadata, accessible for a long time. Once online, information is provided with an adequate version control and is permanently archived.
10. Costs of utilization: By specifying the use of the license: Creative Commons Attribution 3.0 Austria (CC BY 3.0) the user charges are currently not foreseen.

³ Standards und Architectures for E-Government-Applications see http://www.cio.bund.de/DE/Standards/SAGA/saga_node.html

3. Uniform name for administrative management

The Austrian government, like many other European governments, is also influenced by federal principles. Therefore, this section recommends a unified approach to URI names that consistently identifies the data portals and keeps the cost low.

The name and modelling pattern for URL-based identification

The records of the public administration are assigned to the organizations publishing Open Government Data. Each record must be uniquely identifiable by URL.

- The URL for the Austrian portal is: **data.gv.at**. The URL convention for Open Government Data portals is:
data.organisation.gv.at (E.g. data.wien.gv.at; data.statistik.gv.at ...)
- If a more detailed breakdown URL is wished, it is recommended to carry out the subdivision according to the categories defined in the White Paper OGD metadata.
- The URL convention for the database should follow the scheme:
data.organisation.gv.at/katalog
- The URL convention for the terms of use follows the pattern:
data.organisation.gv.at/nutzungsbestimmungen
(See chapter Legal Requirements)
- The URL convention for the formats follows the scheme:
data.organisation.gv.at/formate

It is recommended to publish an accompanying documentation for the data formats. The files should be accompanied by a manual which relates to the format and how it can be used with respect to the data.

Should the name of the organisation change, the old link must be preserved for 5 years at least.

4. Technical requirements - Open Government Data Formats & Metadata

The Metadata Working Group of Cooperation ODG has developed a white paper to OGD metadata structure in Austria. The relevance for INSPIRE was taken into account. As a result, there are a so-called core metadata, additional optional attributes and a vocabulary for metadata catalogue Austria. In the document the recommended formats for OGD are also listed.

5. Organizational Requirement - OGD monitoring

In order to further increase the transparency of data, a data monitoring identifying the data sets for publishing and a process for the publication of data should be set up. The responsible administrative unit can assess internal databases with regard to the following criteria and identifies the records to be released as Open Government Data.

⁴

For the assessment and monitoring of OGD the following catalogue with detailed descriptions of the indicators can be used. Individual criteria evaluated with 0 points are reasons for an exclusion of publication. The sum of the points awarded helps in ranking the first to be processed or records to be published:

Criterion		Review	Pattern Review (0-5)
Non-disclosure/ legal restrictions	Is the data subject to non-disclosure obligations or other legal restrictions or does it include data critical for infrastructure?	0: non-disclosure obligation 1: restrictions exist, are hardly alterable (e. g. EU restrictions) 2: restrictions exist, alterable (e. g. regional or district council with extraordinary resolution with 2/3 majority) 3: restrictions exist, easily alterable (e. g. regional or district council with simple majority) 4: restrictions exist, very easily alterable (e. g. internal rules and practices) 5: no restrictions	5

⁴ KDZ and Vienna, Open Government Implementation Model <http://www.kdz.eu/de/open-government-vorgehensmodell/>

Criterion		Review	Pattern Review (0-5)
Personal or corporate references	Does the data include personal references or can individuals or companies be identified?	0: personal data 1: data cannot be made anonymous, missing approval hardly obtainable 2: data cannot be made anonymous, missing approval obtainable 3: approval obtained (e. g. subsidy data) 4: data can be made anonymous 5: no identification of individuals or companies possible, or no infringement of interests in secrecy deserving protection (sec. 8 Data Protection Act ¹¹)	5
Copyright	Is the administrative agency sole proprietor of data copyright?	0: no copyright: disclosure not possible 1: subject to license fees and approval 2: subject to license fees, approval obtained 3: no license fees, subject to approval 4: no license fees, no approval needed 5: sole possession of copyright ensured	5
Value	How high is the estimated value of disclosure for all target groups?	1: very low value 2: low value 3: medium value 4: high value 5: very high value	5
Effort	How high is the effort of disclosure?	0: unjustifiable cost 1: very high cost 2: high cost 3: medium cost 4: low cost 5: very low	5
Content-related data quality	How high is the estimated data quality? (timeliness, completeness, accurateness, faultiness)	0: data quality unjustifiable 1: data quality very low 2: data quality low 3: data quality medium 4: data quality high 5: data quality very high	3
Technical availability	Available data formats and data sources, open standards: OGD formats, extended Five-Star System ⁵ (see Table 4 and Table 5)	1: data available electronically 2: data available in machine-readable format 3: data available in OGD formats 4: data available with URI / as RDF 5: data available as Linked Data	1
Synergy	Are agencies already making the data/services available for other purposes?	1: already voluntarily published 2: soon to be published voluntarily 3: to be published under alterable regional/national laws 4: already published (e. g. INSPIRE, Environmental Information Directive 2003/4/EC...) under an obligation (e. g. statutory, EU or contractual requirements) 5: to be published (e. g. INSPIRE, Environmental Information Directive 2003/4/EG...) under an obligation that is hard to change (e. g. statutory, EU or contractual requirements)	5
			34

Table 1: Criteria for data monitoring and pattern for evaluation in tabular form (Attention: differs slightly from the German whitepaper v1.1.0 taking into account the newer version of the KDZ paper)

⁵ Berners-Lee at Gov 2.0-Expo in Washington proposing a 5 Star Modell see: <http://inkdroid.org/journal/2010/06/04/the-5-stars-of-open-linked-data/>

6. Legal Requirement – Licensing and Terms of Use

For non-public domain data the Creative Commons CC-BY-AT (version 3.0) ⁶ is to be used for the publication of Open Government Data.

In a nutshell, the rights and obligations of the CC BY license are:

You may reproduce the work or the content, distribute and make available to the public, make modifications and adaptations of the work or content, use the work commercially. You must name the original author / copyright holder in the manner specified by him. The Terms of Use shall be published on the Open Government Data platform, the attribution is to be specified:

The attribution of the **organization** as a holder of rights has to be done in the following manner: "Datenquelle: Organisation - data.organisation.gv.at"

The URL convention follows the scheme:

data.organisation.gv.at/nutzungsbedingungen

Should supplements should be required to license, these can be accommodated in a netiquette.

Best practice - example Vienna: data.wien.gv.at/nutzungsbedingungen

Terms of Use

The city of Vienna publishes data in machine-readable form for further use by the citizen and the economy. The re-use should be simplified by standardized and transparent terms of Use.

Creative Commons License

Open Government Data from Vienna is licensed under a Creative Commons Attribution 3.0 Austria license.

Netiquette

- *The City of Vienna has to be mentioned as the owner in the following way: "Datenquelle: Stadt Wien - data.wien.gv.at"*
- *The citizens should be informed about new applications and services using Open Government Data (OGD) of the city of Vienna. The city of Vienna is entitled to publish information about such applications and services and to use it for reporting.*
- *A publication (link) of the applications and services is permitted to the city of Vienna within the scope of the website data.wien.gv.at.*
- *The city of Vienna calls service providers using OGD of Vienna for their applications and services, to actively inform it where these services and applications can be found. Please send information to open@post.wien.gv.at . Data from Open Government Data Vienna*

⁶ <http://creativecommons.org/licenses/by/3.0/at/>

data.wien.gv.at may not be used for applications or publications supporting criminal, illegal, racist, discriminatory, libellous, pornographic, sexist or homophobic activities or incite such activities.

§ 7 of the Copyright Act specifies that laws, ordinances, official decrees, announcements and decisions as well as works made exclusively or primarily for official use shall not enjoy copyright protection. These works are public domain. Hence, the data is also beyond any licensing. Currently, a new sign- the "Public Domain Mark" - is developed, which will be implemented in the next version of the Austrian Creative Commons adaption for public domain works. From this point a use should be considered for public domain works.