

Sustainability of (Open) Data Portal Infrastructures

Open Data Portal Assessment
Using User-Oriented Metrics



This study has been prepared by the University of Southampton as part of the European Data Portal. The European Data Portal is an initiative of the European Commission, implemented with the support of a consortium led by Capgemini Invent, including Intrasoft International, Fraunhofer Fokus, con.terra, Sogeti, 52North, Time.Lex, the Lisbon Council, and the University of Southampton. The Publications Office of the European Union is responsible for contract management of the European Data Portal.

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

Daniele Rizzi – Policy Officer

Email: daniele.rizzi@ec.europa.eu

European Data Portal

Gianfranco Cecconi, European Data Portal Lead

Email: gianfranco.cecconi@capgemini.com

Esther Huyer

Email: esther.huyer@capgemini.com

Written and reviewed by:

Eric Costa

Email: eric.costa@gmail.com

Elena Simperl

Email: elena.simperl@kcl.ac.uk

Johanna Walker

Email: j.c.walker@soton.ac.uk

Last update: 02.03.2020

www: <https://europeandataportal.eu/>

@: info@europeandataportal.eu

DISCLAIMER

By the European Commission, Directorate-General of Communications Networks, Content and Technology. The information and views set out in this publication are those of the author(s) and do not necessarily reflect the official opinion of the Commission. The Commission does not guarantee the accuracy of the data included in this study. Neither the Commission nor any person acting on the Commission's behalf may be held responsible for the use, which may be made of the information contained therein.

Luxembourg: Publications Office of the European Union, 2020

© European Union, 2020



OA-02-20-166-EN-N

ISBN: 978-92-78-42149-6

doi: 10.2830/51660



The reuse policy of European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Except otherwise noted, the reuse of this document is authorised under a Creative Commons Attribution 4.0 International (CC-BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

Note: this document is part of a series of research reports developed on the topic of “Sustainability of (open) data portal infrastructures”, all of which are available on the European Data Portal at <https://www.europeandataportal.eu/en/impact-studies/studies> .

The series is made of the following reports:

1. A Summary Overview
2. Measuring Use and Impact of portals
3. Developing Microeconomic Indicators Through Open Data Reuse
4. Automated Assessment of Indicators and Metrics
5. Assessment of Funding Options for Open Data Portal Infrastructures
6. Open Data Portal Assessment Using User-Oriented Metrics
7. Leveraging Distributed Version Control Systems to Create Alternative Portals

Executive Summary

The European Data Portal Analytical Report “The Future of Open Data Portals”¹ presented 10 ways in which open data portals can be organised for sustainability and added value:

1. Organising for use of the datasets (rather than simply for publication);
2. Learning from the techniques utilised by recently emerged commercial data marketplaces; promoting use via the sharing of knowledge, co-opting methods common in the open source software community;
3. Investing in discoverability best practices, borrowing from e-commerce;
4. Publishing good quality metadata, to enhance reuse;
5. Adopting standards to ensure interoperability;
6. Co-locating documentation, so that users do not need to be domain experts in order to understand the data;
7. Linking datasets to enhance value;
8. Being measurable, as a way to assess how well they are meeting users’ needs;
9. Co-locating tools, so that a wider range of users and re-users can be engaged with;
10. Being accessible by offering both options for big data, such as Application Programme Interfaces, and options for more manual processing, such as comma separated value files, thus ensuring a wide range of user needs are met.

This report makes these 10 user-oriented sustainability principles fully actionable, by presenting appropriate metrics to assess each principle, and methods by which to measure these. In addition, we test the proposed metrics by assessing them on 10 open data portals.

Table of Contents

Executive Summary	3
1 Introduction	6
2 Methodology	7
3 Metrics	10
3.1 Organise for Use	10
Metric for “Organise for Use”	11
3.2 Promote Use	12
Metric for “Promote Use”	12
3.3 Be Discoverable	13
Metric for “Be Discoverable”	13
3.4 Publish Metadata	14
Metric for “Publish Metadata”	15
3.5 Promote Standards	15
Metric for “Promote Standards”	16
3.6 Co-locate Documentation	17
Metric for “Co-locate Documentation”	17
3.7 Link Data	17
Metric for “Link Data”	18
3.8 Be Measurable	19
Metric for “Be Measurable”	20
3.9 Co-locate Tools	20
Metric for “Co-locate Tools”	22
3.10 Be Accessible	22
Metric for “Be Accessible”	23
4 Assessment	24
4.1 Organise for Use.....	25
4.2 Promote Use.....	25
4.3 Be Discoverable	26
4.4 Publish Metadata	26
4.5 Promote Standards	27
4.6 Co-locate Documentation	28
4.7 Link Data.....	28
4.8 Be Measurable	29
4.9 Co-locate Tools.....	29
4.10 Be Accessible	30
5 Recommendations	31
6 Conclusion	31
7 Referenced Studies	33

Tables

Table 1	5-Level Maturity Schema for Metadata Management	14
Table 2	Metrics Assessment: Organise for Use	25
Table 3	Metrics Assessment - Promote Use	26
Table 4	Metrics Assessment - Be Discoverable	26
Table 5	Metrics Assessment - Publish Metadata	27
Table 6	Metrics Assessment - Promote Standards	27
Table 7	Metrics Assessment - Co-locate Documentation	28
Table 8	Metrics Assessment - Link Data	28
Table 9	Metrics Assessment - Be Measurable	29
Table 10	Metrics Assessment - Co-locate Tools	29
Table 11	Metrics Assessment - Be Accessible	30

Figures

Figure 1	10 Ways Open Data Portals Should Organise for Sustainability and Added Value	7
Figure 2	Open Data Maturity Clustering (EU28, 2018))	8
Figure 3	Measuring UX Website	10
Figure 4	Web Quality Checklist for Open Data Sites (Opquast)	11
Figure 5	Metadata Structure for Open Data Portals	14
Figure 6	DCAT Application Profile for Data Portals in Europe	16
Figure 7	5-Stars Scheme for Linked Open Data	18
Figure 8	Google Analytics	20
Figure 9	Visualisation Tools of the EU Open Data Portal	21
Figure 10	Visualisation Tools of Eurostat	21
Figure 11	Automated Accessibility-Checking Tools Suggested by Tullis and Albert	23

1 Introduction

Open data portals are central points of access for datasets.¹ Accordingly, there is a need for new initiatives to define metrics to assess the quality of open data portals.² In their 2015 report on open data maturity in Europe, the European Data Portal also mention that only a very few countries have evaluated their open data strategies ex-post, so there is a need to measure the success of current open data initiatives. This can be achieved by defining a number of key performance indicators and benchmarks that can be measured over time, and against other portals.³

Looking at previous projects and research, different approaches have been developed for open data and open data portals evaluation:

- Global Open Data Index and Open Data Census developed by the Open Knowledge Foundation:^{4,5}
 - track the status of open government data (OGD), i.e. which countries are publishing data in a timely and right way;
 - compare the progress made by different cities and local areas in releasing open data.
- Open Data Barometer developed by the World Wide Web:⁶
 - focus on open data readiness, implementation, and emerging impacts.
- Open Data Maturity Model developed by the Open Data Institute:⁷
 - assess how well an organisation publishes and consumes open data and identify actions for improvement.
- Portal Data Quality Assessment project:⁸
 - evaluate the quality of a Web Portal by defining a data quality model.
- Web Quality Checklist for Open Data Sites developed by Opquast:⁹
 - provide a checklist for open data publishing, including some quality aspects.

By reviewing existing open data assessment frameworks, Welle Donker and van Loenen revealed that these only covered part of the open data ecosystem. Therefore, they developed a framework to assess the maturity of the open data ecosystem by assessing open data supply, open data governance, and open data user characteristics holistically.¹⁰

Another relevant work in this area is that developed by Colpaert et al., where the authors introduce the 5 stars of Open Data Portals. This framework serves as a guide for Public Administrations that are new to the open data field, as well as for existing and established open data portals, to help them work progressively towards more involved and advanced goals:¹¹

- ★ A dataset registry: a list of links to Open Datasets.
- ★★ A meta-data provider: maintain, structure and open up your meta-data.
- ★★★ A co-creation platform: gather tools and stimulate conversations on re-use.
- ★★★★ A data publishing platform: provide the data itself in common formats.
- ★★★★★ A common data hub: open governance, provenance, trust and versioning.

There are also a number of approaches to measuring the quality of the data and metadata in the portal, such as the work of Umbrich et al.

However, it is insufficient simply to publish good quality data to a portal, as it risks creating “virtuous data dumps”.¹ It is equally important to focus on other aspects of publishing, managing and using data, by understanding the needs of average citizens and data professionals and by choosing adequate tools to deliver the capabilities and user experience that they require. Analytical Report 8 (AR8) presented 10 ways in which open data portals must evolve for sustainability and added value (**Error! Reference source not found.**).

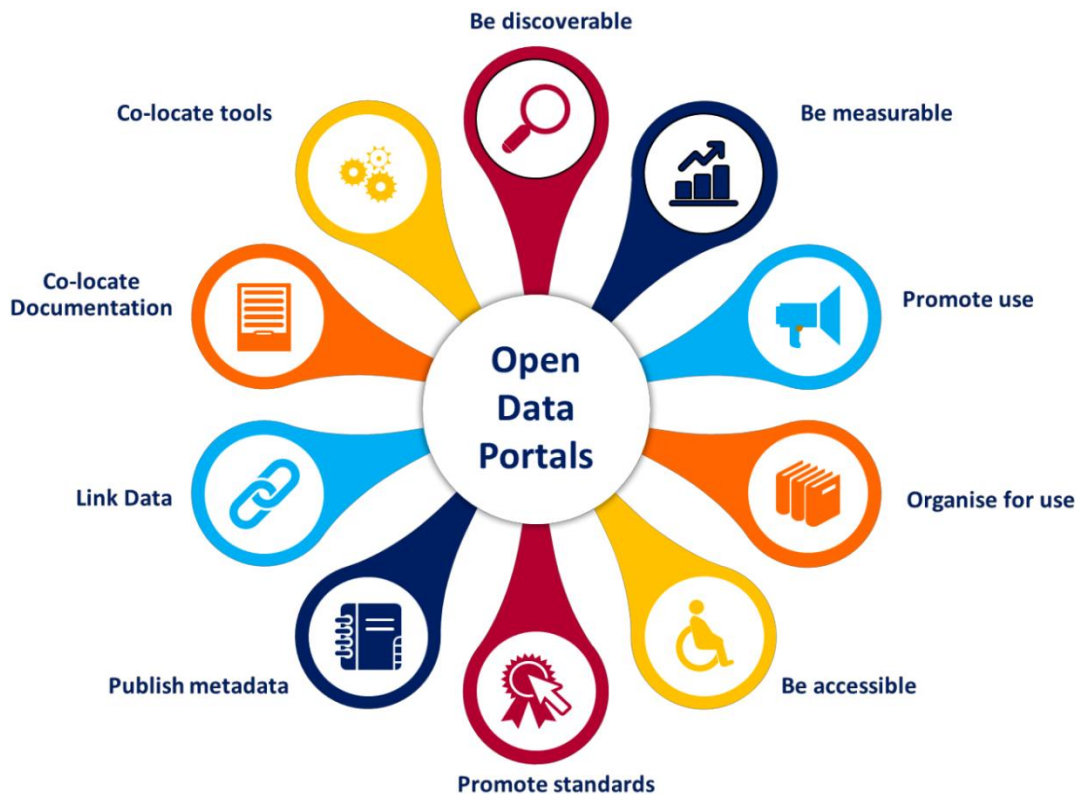


Figure 1 10 Ways Open Data Portals Should Organise for Sustainability and Added Value

The main aim of this report is to provide portal owners with metrics and guidelines so they can know more about their portals, achieve active use and impact and improve their user experience.

2 Methodology

Following these 10 user-oriented sustainability principles, this report presents appropriate methods and metrics to assess each principle. To obtain our metrics, we used a variety of information sources, including published academic papers, white papers, independent reports and initiatives from the European Commission and from other recognised institutional entities. Google Search, Google Scholar, Mendeley, Research Gate and Web of Science were the search engines and bibliographic

¹ Leonard, S., (2012). The Fog of More. *The New Inquiry* <https://thenewinquiry.com/the-fog-of-more/>

databases used to obtain the list of references for this report, as well as using cross referencing. Using the proposed metrics, we assessed the conformity of 10 open data portals to these user-oriented sustainability principles. We selected portals from different stages of Open Data Maturity, retrieved from the 2018 report “Open Data Maturity in Europe” of the European Data Portal (**Error! Reference source not found.**):¹²

- **Beginners:** Think big, act small;
- **Followers:** Strengthen governance, boost engagement;
- **Fast-trackers:** Graduate from traction to impact;
- **Trend-setters:** Maintain the ecosystem, experiment and share the knowledge.

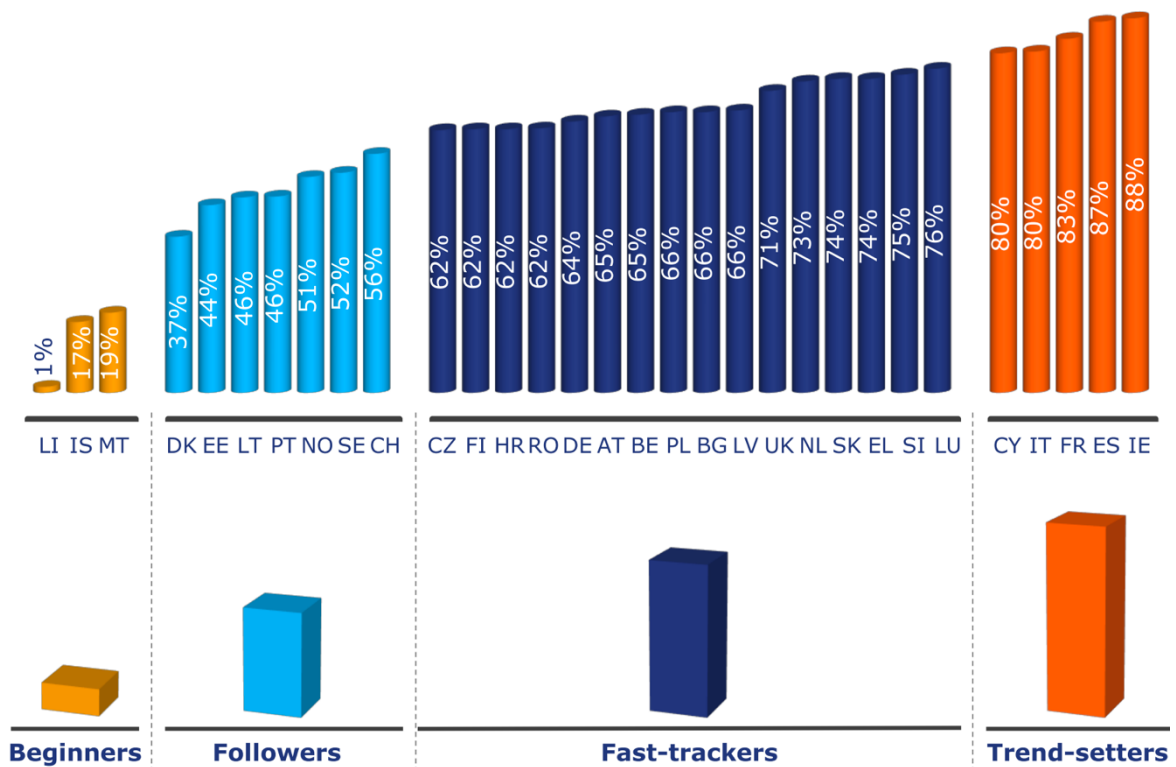


Figure 2 Open Data Maturity Clustering (EU28, 2018))

In addition to EU government data portals, we also selected some other specific open data initiatives that we found to be interesting to include in this study. Thus, the 10 open data portals assessed in this report are:

- Cyprus National Data Portal (trend-setter);¹³
- Avoindata.fi (fast-tracker);¹⁴
- Data.gov Belgium (fast-tracker);¹⁵
- Data.gov Slovakia (fast-tracker);¹⁶
- Dados.gov Portugal (follower);¹⁷
- Island.is (beginner);¹⁸
- +
- EU Open Data Portal;¹⁹
- London Datastore;²⁰

- Geo Data Portal Luxembourg;²¹
- Open Data Trento.²²

3 Metrics

3.1 Organise for Use



According to AR8, open data portals are often not organised with the user experience in mind. The authors argue that to offer an added value for users, portals need to be organised for use of the datasets, rather than simply for publication. A crucial element in this context is to analyse user behaviour and user experience (UX) to ensure that portals are meeting user needs.

Many researchers and practitioners have been using different types of UX methods and tools for measuring the user experience. Tullis and Albert provide a detailed practical guide on collecting, analysing, and presenting a wide variety of usability metrics (performance metrics, issue-based metrics, self-reported metrics, behavioural and physiological metrics, and combined and comparative metrics).²³ This work is supported by a specific website Measuring the User Experience²⁴ (Figure 3) that provides some resources, including publications, blogs, spreadsheets, tools and services.

Measuring the User Experience
Collecting, Analyzing, and Presenting Usability Metrics
A companion website to the book by Tom Tullis and Bill Albert

- About the Book
- About the Authors

Publications & Presentations

- 2D vs. 3D Charts: Does 3D Representation Help or Hurt? by Tom Tullis and Michaela Case, UXPA 2017 International Conference, Toronto (PDF of Poster) (June 2017)
- Introducing a New UX Maturity Metric: Team Engagement Score (TES) During Usability Testing by Bill Albert and Josh Rosenberg, UXPA International, Toronto (June 2017)
- Bill gave a workshop on Measuring Emotional UX Engagement at UX Hong Kong (February 2017)
- Bill gave a keynote presentation on Moving your UX Career to the Next Level at UX Hong Kong (February 2017)
- What's the Most Effective Way to Present Two-State Toggle Buttons? by Tom Tullis, Azilah Baker, and Lori LeDoux, UXPA 2016 Conference, Poster Session, Seattle, Washington (PDF of paper) (June 2016)
- Does Ad Blindness on the Web Vary by Age and Gender? by Tom Tullis and Marisa Siegel, ACM CHI 2013 Conference on Human Factors in Computing Systems (PDF of paper) (April 2013)
- Can Users Remember Their Pictorial Passwords Six Years Later? by Thomas S. Tullis, Donna P. Tedesco, & Kate E. McCaffrey, ACM CHI 2011 Conference on Human Factors in Computing Systems (PDF of paper) (May 2011)
- Did You See That Thing? An eye tracking study on the reliability of self-reported awareness measures by Bill Albert and Donna Tedesco. UPA-Boston's Ninth Annual Mini UPA Conference (June 9, 2010)
- What's the Latest: Research on the Design of Web Pages Show? by Tom Tullis, Fiona Tranquada, and Marisa Siegel. UPA-Boston's Ninth Annual Mini UPA Conference (June 9, 2010)
- Squeezing 1,000 Users into the Lab, Or, How to Conduct Online User Experience Studies by Bill Albert, Donna Tedesco, and Tom Tullis. Presentation to UPA Boston (May 18, 2010)
- Webinar: Combining Lab and Online Usability Testing: Lessons Learned by Bill Albert and UserZoom (May 6, 2010).
- Debunking the Myths of Online Usability Testing by Bill Albert, on JohnnyHolland.org (April 9, 2010).
- Reliability of Self-Reported Awareness Measures Based on Eye Tracking by William Albert and Donna Tedesco. *Journal of Usability Studies* (February 2010).
- Pre-Conscious Assessment of Trust: A Case Study of Financial and Health Care Web Sites by William Albert, William Gribbons, & Jindrich Almadras. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting—2009*.
- Top Ten Myths About Usability: Keynote Presentation by Tom Tullis at World Usability Day, Simmons College, Boston, MA (November 12, 2009)
- Usability Marathon: Beyond the Usability Lab: Exploring Large-Scale User Experience Research by Bill Albert and Lena Dmitrieva (October 12, 2009)
- Are People Drawn to Faces on Webpages? Video of Five-minute Presentation by Tom Tullis at Ignite Boston 6 (September 17, 2009) (based on CHI paper)
- Rating Scores: What the Research Says by Joe Dumas and Tom Tullis (May 2009)

Announcements

- Check out our new page of System Usability Scale (SUS) Resources.
- Bill and Tom are starting on the 3rd Edition of *Measuring the User Experience!* Do you have the book? We would appreciate your feedback about it and suggestions for improvement. Take this short (5-min) survey.
- Check out Tom's new website, UX Metrics Geek, and his UX Trivia Quizzes.
- Bill was interviewed by Cranberry Radio on UX at the Conversion Conference, June 2016.
- Tom was awarded the 2011 Lifetime Achievement Award by the Usability Professionals Association at their 20th Anniversary Conference.

Articles & Blogs

- Five Takeaways from the UXPA 2018 Conference by Tom Tullis (July 6, 2018)
- Five Tips for a Successful Career as a UX Researcher by Tom Tullis (March 21, 2018)
- Five Tips for Using Surveys in UX Research by Tom Tullis (Nov. 10, 2017)
- Worst Usability Issue? by Tom Tullis (July 4, 2011)
- Results of Online Usability Study Comparing Obama and McCain Websites (Nov. 14, 2008)
- A Timestamp Macro for Task Timing (March 29, 2008)
- Calculating a Confidence Interval for Task Completion using the Adjusted Wald Method (March 28, 2008)
- Results of Online Usability Study of Apollo Program Websites (March 21, 2008)
- OK and Cancel Buttons: What's the Right Order? (Feb. 28, 2008)

Consulting Services

- The User Experience Center at Bentley University provides consulting services to corporate clients worldwide. Bill Albert, the Director of the center, heads up a team of full-time UX professionals conducting a wide variety of user research, including quantitative benchmarking, unmoderated testing, eye tracking, and usability testing. Contact Bill for an estimate: walbert@bentley.edu or 1-781-891-2608.

Tools and Services

- System Usability Scale (SUS) Resources
- Card-sorting Tools
- Online Survey Tools

Figure 3 Measuring UX Website

Nevertheless, as one might expect, the metrics and resources provided are focused only on assessing the behaviour and obtaining feedback from users. Therefore, we had to find new metrics specifically to enable portal owners to organise their portals following a user-centric approach. What we desire in the first instance is for portal owners to self-evaluate their open data portals and subsequently take action for improvement. After that, in future extensions of this work, i.e. after the implementation of the required improvement actions, we might use or recommend some UX methods and tools to measure the actual user experience.

In order to define the metrics to organise portals for use of the datasets, we used some items of the “Web quality checklist for open data sites”⁹, developed by Opquast²⁵, a French e-learning digital intelligence company specialising in improving the digital ecosystem to efficiently enhance productivity and optimise user experience.

Opquast’s checklist for open data sites includes 72 items, divided into 13 themes and 3 levels (e.g. level 1 is for basic requirements and level 3 practices stand for higher applicability) (Figure 4). This checklist has been cited and used by previous research.^{2,26}

The screenshot shows the Opquast OpenData checklist interface. At the top, the Opquast logo and 'Check-lists' are displayed. Below the logo, there is a navigation bar with 'Accueil' and 'OpenData'. The main content area is titled 'OpenData' and contains a table with 7 items. To the right, there is a promotional graphic for a book titled 'QUALITE WEB'.

N°	Rubrique	Description
		Faire une recherche
1	Animation	The concept of Open Data is explained
2	Animation	The site provides a way for users to keep informed of updates to the data
3	Animation	The site provides an online channel for feedback
4	Animation	Examples of reuse (fictitious or real) are provided
5	Animation	A news feed provides information about updates to each dataset
6	Animation	Datasets are accompanied by links or resources that provide user guidance and support
7	Animation	There is a discussion forum about the datasets

Figure 4 Web Quality Checklist for Open Data Sites (Opquast)

Metric for “Organise for Use”

Accordingly, we constructed a metric based on previous research and on the theme of usability of the Opquast’s checklist, adapted for the case of portals. For any given open data portal, one point is awarded for each of the following items:

- Each dataset is accompanied by a comprehensive descriptive record (going beyond a collection of structured metadata).
- An extract of the data can be previewed (for easier sense making).
- The portal provides recommendations for related datasets.
- The portal enables users to review/rate the datasets.
- Keywords from datasets are linked to other published datasets.

3.2 Promote Use



Another user-oriented principle suggested in AR8 is to promote use of the datasets on the portals, through the sharing of knowledge and co-opting methods. These techniques are common in the open source software community. Other tools and techniques suggested are impact stories, examples, curated list of datasets, useful lists, rating systems, discussion forums, Q&A, and other social channels.

As a way to integrate two emerging government trends of using Web 2.0 social media and open data, Alexopoulos et al. develop what they call a “*second generation of open government data platform*”, based on Web 2.0 features: a set of capabilities for data processing; enhanced data modelling (flat, contextual and detailed metadata); commenting existing datasets and expressing needs for new datasets (feedback and collaboration); datasets quality rating; users groups formation and extensive communication and collaboration among them; data linking; publication and upload of new versions of existing datasets; and advanced data visualisation.²⁷

Máchová and Lněnička propose and validate a benchmarking framework to evaluate the quality of open data portals on the national level. Some communication and participation features are also included in their framework:²⁸

- **Forum:** to submit some user feedback on the datasets and to discuss and exchange ideas.
- **Request form:** to request or suggest a new format type of open data.
- **Help functionality:** to learn how to use the portal and improve the usability.
- **FAQ:** to help resolve any recurrent issues.
- **Connection to social media:** to create a social distribution channel for open data (e.g. users can share how they used and learned from a dataset).

Goh et al. explore some knowledge management mechanisms to support the access, creation and transfer of knowledge between portals and their users. The dimensions considered for the knowledge transfer, i.e. to develop a wider use of data, are the following:²⁹

- Online collaboration:
 - Organisation-to-user collaboration (e.g. ask-an-expert feature);
 - User-to-user collaboration (e.g. discussion forums, blogs, wikis, instant messaging).
- Information alerts (e.g. newsletters, events calendar, update frequency of documents, RSS feed, mobile alerts);
- User support (e.g. FAQs, helpdesk, search tips, tutorials, demos);
- Resource sharing (e.g. own repository of information, links to other websites, information contributed by other users).

Metric for “Promote Use”

By considering these previous studies together with some items of the Opquast’s checklist,⁹ we constructed our metric to assess the user-oriented principle of promoting use of datasets. For any given open data portal, one point is awarded for each of the following items:

- The portal is connected with social media to create a social distribution channel for open data.
- The portal provides users with online support for feedback, to request/suggest the publication of new datasets, and when problems arise during use (e.g. contact form, discussion forum, FAQs, helpdesk, search tips, tutorials, demos).
- The portal provides a way for users to keep informed of updates to the data (e.g. news feed).
- Datasets are accompanied by links or resources that provide user guidance and support.
- Examples of reuse (fictitious or real) are provided (e.g. information contributed by other users, last reuse, best reuse, data stories).

3.3 Be Discoverable



AR8 states that to create an added value in open data portals it is vital to invest in discoverability best practices. Borrowing from e-commerce, to facilitate advanced searches, data portals should be enabled for search engines, instead of attempting to engage users with Boolean logic. Another suggestion is identifying a dataset as existing but not available, to offer greater transparency, to limit the time wasted on abortive searches, and to show visitors that the publisher is monitoring demand.

According to Attard et al., one major challenge in open government data initiatives is the discoverability. The discoverability of open data is associated to the quality of the metadata, which is not always complete or accurate. Another problem is that some portals only have simple search functions, which results on users to spend more time on finding relevant datasets. Accordingly, the proposed solution from these authors is to use good quality metadata and more advanced search tools on portals to improve discoverability.³⁰

Metric for “Be Discoverable”

For this user-centred principle we use the discoverability metric developed by Walker et al., as this metric has already been studied and developed by these authors. They focused on an approach based closely on problems faced in discovering appropriate data, so they constructed a metric based on the availability of solutions to these challenges. Thus, for any given dataset, one point is awarded for each of the following items:³¹

- The publisher/owner of the data has an open data portal (or similar search mechanism).
- The publisher/owner of that portal publishes an updated, searchable list of datasets.
- The publisher/owner of that portal publishes an updated, searchable list of datasets with synonyms.
- The publisher/owner of that portal publishes a list of datasets which are known to exist but are not currently available (limiting the time wasted on abortive searches).

3.4 Publish Metadata



AR8 recommends that publishing good quality metadata is fundamental to enhance reuse, findability and cataloguing, as well as to make associations and relationships between datasets.

Metadata is a well-researched area and many methods and approaches have been proposed for the production, measurement and analysis of metadata. Figure 5 presents a summary of the most common and most important metadata of open data portals.^{28,32,33}

title	description	publisher	frequency	release date
update date	temporal coverage	geographic coverage	license	data dictionary
granularity	metadata update	dataset URL	dataset size	dataset format
number of views	number of downloads	user rating	dataset theme	tags/keywords

Figure 5 Metadata Structure for Open Data Portals

In 2011, following a similar approach to Tim Berners-Lee’s 5-star Open Data schema,³⁴ the European Commission created a 5-level maturity schema for metadata management, to help Member States to identify directions for improving their own metadata management policies (Table 1).³⁵

Table 1 5-Level Maturity Schema for Metadata Management

Level	Description	Direction/recommendation
1 - Metadata Ignorance	Metadata is not documented, mainly because administrators are not aware of its importance. This situation results in serious semantic interoperability problems within each country as developers use ad hoc data models, metadata, code-lists, taxonomies, etc for developing eGovernment systems which are very hard and expensive to interoperate later due to competing data specifications.	Public administrations should become aware of the importance of Metadata in eGovernment and the need for coherent relevant management policies.
2 - Scattered or Closed Metadata	Metadata may be partially documented but a) not in a centralised and structured way or b) it is not available and accessible under an open license	Public administrations should organize the scattered Metadata in structured repositories, catalogues or libraries and

	framework, in other words as "Open Metadata" for developers to share and reuse.	provide open access to the collected resources.
3 - Open Metadata for Humans	Metadata is documented and becomes available as "Open Metadata" for reuse, but are not systematically published in a reusable format, e.g. may only be available in .pdf or .doc documents.	Public administrations should provide services to query, browse and export their Metadata in a machine-readable and preferably non-proprietary format (e.g. CSV, XML).
4 - Open Reusable Metadata	Metadata is centrally managed, and published as "Open Metadata", in a machine-readable format and/or an API is provided for computers to access, query and reuse the available metadata repositories, catalogues, libraries, etc. Electronic Metadata Management Systems are introduced to support metadata architectures and policies. Through these systems users can find, browse, compare, download and use Metadata that better fits their needs and projects.	Public administrations should consider applying linked metadata policies, including use of Resource Description Framework (RDF) to document their Metadata, persistent design, use and maintenance of Universal Resource Identifiers (URIs), linking to external vocabularies/schemata, harmonize their resources to third parties' resources etc.
5 - Linked Open Metadata	Semantic Assets are documented using linked data principles and are managed by advanced Metadata Management Systems. At this level, a graph of interlinked concepts emerges, as Metadata definitions systematically use definitions from other vocabularies. Each concept represented in this graph corresponds to a unique URI.	Public administrations should consider applying linked metadata policies, including use of RDF to document their Metadata, persistent design, use and maintenance of URIs, linking to external vocabularies/schemata, harmonize their resources to third parties' resources etc.

Metric for "Publish Metadata"

For the user-oriented principle of publishing metadata, we propose using the 5-level maturity schema presented in Table 1:

- ★ Metadata Ignorance.
- ★★ Scattered or Closed Metadata.
- ★★★ Open Metadata for Humans.
- ★★★★ Open Reusable Metadata.
- ★★★★★ Linked Open Metadata.

3.5 Promote Standards



In the case of datasets standards, AR8 suggests that adopting standards is important to ensure interoperability.¹ For example, using the Data Catalogue Vocabulary Application Profile for data

portals in Europe (DCAT-AP) (Figure 6),³⁶ common metadata standards are applied across multiple data portals, which can subsequently enable a cross-data portal search for datasets. This can be accomplished by the exchange of descriptions of datasets among data portals.

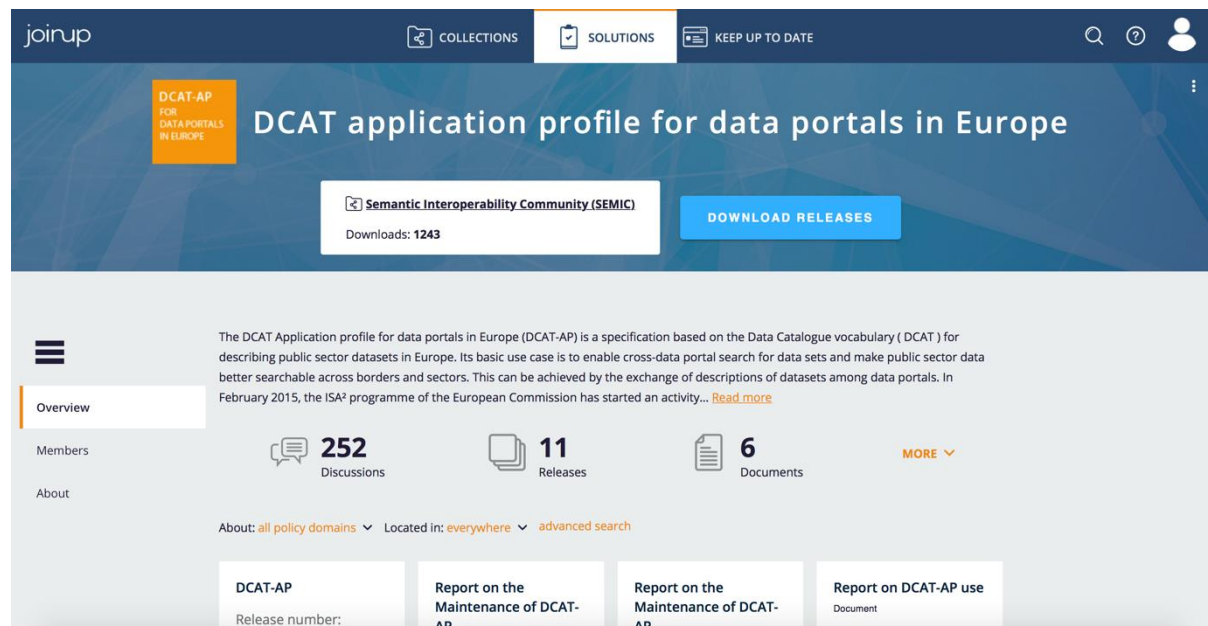


Figure 6 DCAT Application Profile for Data Portals in Europe

Based on Tim Berners-Lee's 5-stars scheme for Linked Data,³⁴ the Open Data Institute's Open Data Certificates allows to assess and certify data portals that meet standards for publishing sustainable and reusable data. If we look at the 4th star of this same 5-stars scheme, using standards becomes crucial as it is recommended to "use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff".

Metric for "Promote Standards"

In order to promote datasets standards, we constructed a metric based on guidelines from the W3C eGov Interest Group³⁷ and on some items selected from the Opquast checklist⁹. Once again, one point is awarded for each of the following items:

- A permanent, patterned and/or discoverable URI/URLs is used for each dataset (e.g. URI/URLSs can be used as universal, unique identifiers by appending a serial number or other internal naming system to a domain).
- The portal uses versioning of datasets (to maintain the history of a dataset).
- Dates are available in a standard format (facilitates the automated exploitation of date-type data and their conversion according to specific needs or constraints).
- Metadata associated with each dataset is available in a standard format (e.g. using VOID or DCAT) to enable automated metadata retrieval and import of metadata from other data catalogues.
- The metadata catalogue can be retrieved using a standard protocol (e.g. automatic retrieval of the metadata catalogue using RDF or HTTP GET).

3.6 Co-locate Documentation



Co-locating documentation is another fundamental aspect required to promote the sustainability of open data portals, so that users do not need to be domain experts in order to understand the data. The authors of AR8 suggest that supporting documentation should be immediately and easily accessed from the dataset and should be context-sensitive, to allow users a direct access to information about a specific item of concern.

Other previous research also demonstrates the importance of supporting documentation in open data portals. Attard et al. explain that documentation helps the data to be more understandable and less ambiguous, allowing an easier data discovery.³⁰ One of the attributes defined in the Portal Data Quality Model of Calero et al. is documentation, which is defined as the amount and usefulness of documents with meta information, i.e. “if data have useful documents with meta information then they will be understood better”.⁸ Finally, Máchová et al. also mention that it is essential for open data portals to offer proper documentation and guidelines to support users in the reuse of open data.³³

Metric for “Co-locate Documentation”

For this user-centred principle we use the intelligibility metric developed by Walker et al. This metric is focused on measuring the availability of supporting information. The metric is the following (with increasing value):³¹

1. Supporting documentation does not exist.
2. Supporting documentation exists, but as a document which has to be found separately from the data.
3. Supporting documentation is found at the same time as the data (e.g. the link to the document is next to the link to the data in the search).
4. Supporting documentation can be immediately accessed from within the dataset but it is not context sensitive (e.g. a link to the documentation or text contained within the dataset).
5. Supporting documentation can be immediately accessed from within the dataset and it is context sensitive so that users can immediately access information about a specific item of concern (e.g. a link to a specific point in the documentation or the text contained within the dataset).

3.7 Link Data



One of the most relevant principles mentioned in AR8 to add value is linking datasets. Successful exploitation of datasets should be achieved with this ability for portals to create links to core reference data. The advantages of linking datasets are to allow the cross-referencing and analysis of multiple datasets to point to previous versions of the same dataset, external datasets not hosted on the portal or recommendations based on content or user features.

We already mentioned in previous sections Tim Berners-Lee’s 5-stars scheme for Linked Open Data (Figure 7).³⁴

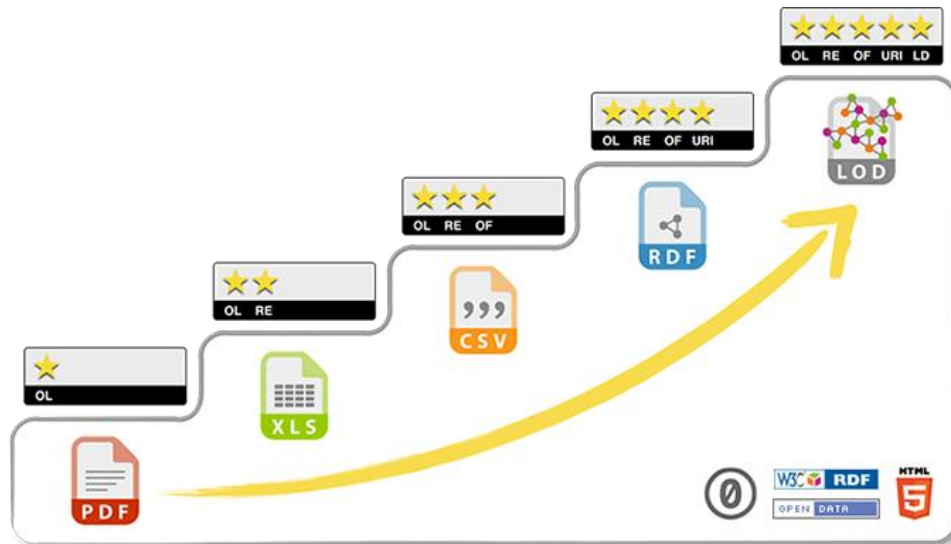


Figure 7 5-Stars Scheme for Linked Open Data

This is the most well-recognised measure of open data, an accepted and easily applicable measure of open data format standards which closely matches the user need to be able to discover unanticipated relationships among data.³¹

Metric for “Link Data”

The metric used for measuring links between datasets is:³⁴

- ★ **On the Web:** Make your stuff available on the Web (whatever format) under an open license.
- ★★ **Machine-readable data:** Make it available as structured data (e.g. Excel instead of image scan of a table).
- ★★★ **Non-proprietary format:** Make it available in a non-proprietary open format (e.g. CSV instead of Excel).
- ★★★★ **RDF standards:** Use URIs to denote things, so that people can point at your stuff.
- ★★★★★ **Linked RDF:** Link your data to other data to provide context.

3.8 Be Measurable



According to AR8, open data portals should be measurable to assess how well they are meeting users' needs. On the one hand, absolute metrics such as completeness, correctness and timeliness could be used to point people to areas in a dataset that require improvement. On the other hand, relative metrics such as fitness of use could increase the confidence in a dataset and act as a differentiator for the selection of a particular dataset. In addition, the inclusion of user reviews is also a good way to measure the usability of portals and the use of digital object identifiers is important to inform users on which data might be considered high value. Nonetheless, a fundamental issue for portals is to meet the needs of different user groups, e.g. in some scenarios, a dataset with low timeliness can be relevant for a particular user group, so it should not be ignored or de-emphasised.

Research in the area of E-Commerce indicates that there are several tools to analyse the usability of E-Commerce websites, such as embedded tools in the systems for deducing statistical information, web analytics free tools (e.g. Figure 8, Google Analytics), and tools for extracting useful knowledge from web resources, i.e. web mining (e.g. RapidMiner and R). Some metrics are also proposed, which can be classified in two major groups:³⁸

- Common indicators:
 - Number of visits;
 - Unique visitors;
 - Repeat visits;
 - Duration of site visits;
 - Exit rate;
 - Number of pages, viewed by one visitor;
 - The most popular pages;
 - Exit pages;
 - Countries of registered visits.
- Specific indicators:
 - Average rate of a visit;
 - Rate of a visit respective to clients' types;
 - Rate of visits respective of countries;
 - Clients' loyalty;
 - Reviewed goods, which are a good source of profits;
 - Successions of connected events and goods, which are purchased together;
 - Exceptions and risk control;
 - Effectiveness of advertising campaigns.

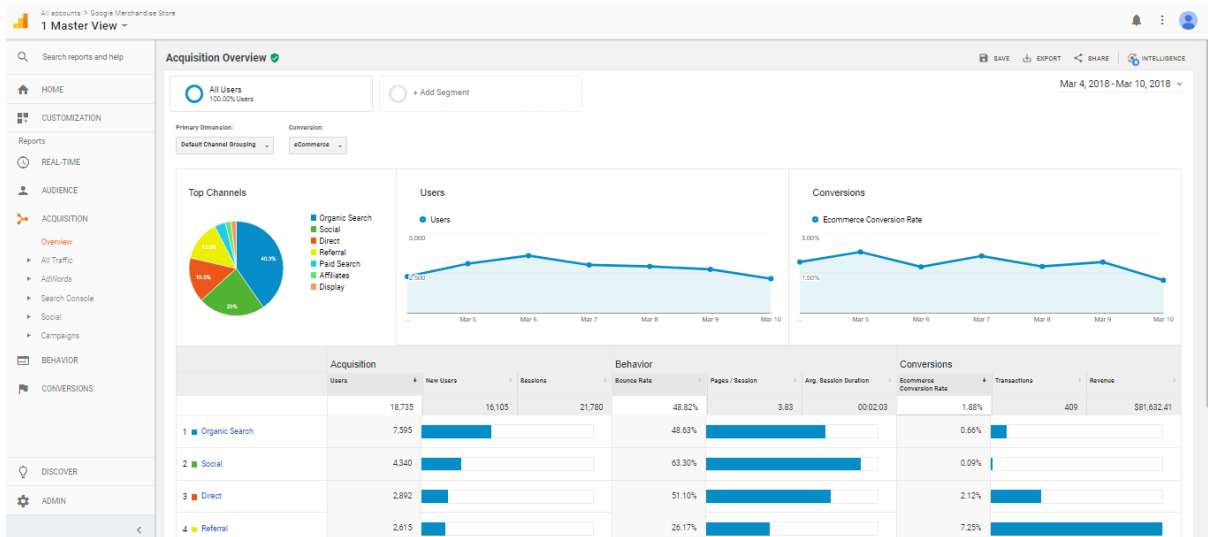


Figure 8 Google Analytics

Metric for “Be Measurable”

A possible scale for the user-centred principle of “Be Measurable” might be (with increasing value):

1. Portal has No analytics.
2. Portal has Site analytics.
3. Portal has Use analytics.
4. Portal has Impact analytics.

3.9 Co-locate Tools



Another relevant user-oriented principle proposed in AR8 is the ability to co-locate tools so that a wider range of users and re-users can be engaged with the datasets of an open data portal. Mapping and visualisation tools can have a huge impact on an individual’s ability to explore a dataset and decide on its relevance.

By evaluating the usability of open data portals, Máchová et al. also stress the importance for portals to extend their features with advanced search capabilities and, visualisation and analytics tools, as most portals only allow users to just download the available data.³³

Many other previous studies also suggest the adoption of data visualisation interfaces (visualisation in charts, maps, plots, etc.) to increase the levels of re-use, as well as to work with multiple, or even across, data catalogues.^{27,28,32}

Some examples of best practice are the visualisation tools provided by the EU Open Data Portal (Figure 9)³⁹ and Eurostat (Figure 10)⁴⁰.

EU Open Data Portal
Access to European Union open data

EUROPA > EU Open Data Portal > Visualisations - β Share


Home Data Applications Linked data Visualisations - β Developers' corner About

Overview Tools Knowledge center Projects Sandbox

Visualisation catalogue

The catalogue is a collection of visualisation tools, trainings and re-usable visualisations for all levels of data visualisation expertise, from beginner to expert.

Total number of entries: 102



Tools

[view all >](#)

Leaflet.js
Open-source JavaScript library for mobile-friendly interactive maps.
Input Formats: [GEOJSON](#)

GeoJSON Editor (GEOD)
The GeoJSON Editor (GEOD) is a map editor allowing you to draw shapes (lines, polygons) and create markers on a map, in a multilingual way. Features can be exported an...
Input Formats: [GEOJSON](#) | [CSV](#) | [ZIP](#)
Output formats: [GEOJSON](#) | [CSV](#) | [ZIP](#)

Basic Map (BMAP)
Basic map (BMAP) is a "light" version of the Commission Map service. It allows configuring a simple, static map without the need to define business logic in ...
Input Formats: [JSON](#)
Output formats: [INTERACTIVE MAPS](#)

Knowledge center

[view all >](#)

Introductory training on Data Visualization
Introductory Data Visualisation training, created by the Publications Office of the European Commission. It contains tools demonstrations in several data visualisation...
Type: Training
Related tools: [D3.js](#), [Google-Visualization](#), [Raw](#), [Tableau](#), [Qlik Sense](#)

Browse tools by visualisation type

- Bars
- Dots / bubbles
- Grids
- Infographics
- Lines
- Maps
- Relations/dependencies
- Shapes/proportions
- Other

Browse tools by data type

- Geospatial data
- Network data
- Numerical data
- Textual data
- Other

Figure 9 Visualisation Tools of the EU Open Data Portal

VISUALISATIONS, MOBILE APPS & EXTRACTION TOOLS

In recent years, Eurostat has developed a variety of data visualisation tools in order to better meet the needs of our users. These tools present data from different statistical themes in an attractive and easy-understandable way for everyone to explore.

On this page you will find an overview of all Eurostat data visualisation tools as well as mobile apps and tools offered for data extraction.

VISUALISATION TOOLS



THEMES IN THE SPOTLIGHT

ECONOMIC TRENDS

MY CAPITAL IN A BUBBLE

MY COUNTRY IN A BUBBLE

GOVERNMENT EXPENDITURE

YOU IN THE EU

QUALITY OF LIFE

YOUNG EUROPEANS

LIVES OF THE ELDERLY

QUIZ

Figure 10 Visualisation Tools of Eurostat

Finally, Sayogo et al. propose a scale to assess the level of participation and collaboration of users by looking at existing tools and/or systems of open government data (OGD) portals:⁴¹

- **No features:** The OGD portal does not provide tools and/or systems for users to engage with government and/or other users.
- **Participative:** The OGD portal provides tools and systems to enable users to participate in the governance of the portal such as voting or rating but the engagement with other users is limited or mediated by the administrator.
- **Collaborative:** The OGD portal provides tools, methods and systems to enable users to innovatively and collaborate with other users be it engage government at any level, non-profit organizations, businesses or other enthusiasts.

Metric for “Co-locate Tools”

Our metric for this principle was constructed based on previous studies and on some items of the Opquast checklist.⁹ A possible scale might be (with increasing value):

1. The portal does not provide visualisation or collaboration tools for users to engage with the datasets.
2. The portal provides visualisation tools to enable users to engage with the datasets.
3. The portal provides visualisation and collaborations tools to enable users to participate in the governance of the portal (e.g. dataset rating) but the engagement with other users is limited or mediated by the administrator.
4. The portal provides visualisation and collaborations tools to enable users to collaborate innovatively with other users.

3.10 Be Accessible



According to AR8, portal owners should make an additional effort to work with data publishers to improve publication formats, making datasets more accessible to users. This could help to solve the problem of having nearly one quarter of the datasets in open data portals published as non-machine-readable portable document format (PDF).

One of the main principles of Open Government Data is accessibility: “*Data is available to the widest range of users for the widest range of purposes*”.⁴² In 2014, Tauberer extended this principle by adding that data should be made available in formats that support both intended and unintended uses of the data, by being published with current industry standard protocols and formats, preferably open, non-proprietary protocols and formats. Data should be discoverable and be provided with sufficient metadata and documentation so that the user understands the structure of the data.⁴³

One of the most widely recognised web accessibility guidelines are the Web Content Accessibility Guidelines (WCAG), Version 2.0, from the World-Wide Web Consortium (W3C).⁴⁴

- Perceivable
 - Provide text alternatives for non-text content.
 - Provide captions and other alternatives for multimedia.
 - Create content that can be presented in different ways, including assistive technologies, without losing meaning.
 - Make it easier for users to see and hear content.
- Operable
 - Make all functionality available from a keyboard.
 - Give users enough time to read and use content.
 - Do not use content that causes seizures.
 - Help users navigate and find content.
- Understandable
 - Make text readable and understandable.
 - Make content appear and operate in predictable ways.
 - Help users avoid and correct mistakes.
- Robust
 - Maximize compatibility with current and future user tools.

Tullis and Albert suggest some automated tools that can check some violations of these guidelines (Figure 11):²³

AUTOMATED ACCESSIBILITY-CHECKING TOOLS

Some of the tools available for checking web pages for accessibility errors include the following:

- Cynthia Says (<http://www.contentquality.com/>)
- Accessibility Valet Demonstrator (<http://valet.webthing.com/access/url.html>)
- WebAIM's WAVE tool (<http://wave.webaim.org/>)
- University of Toronto Web Accessibility Checker (<http://achecker.ca/checker/>)
- TAW Web Accessibility Test (<http://www.webdevstuff.com/103/taw-web-accessibility-test.html>)

Figure 11 Automated Accessibility-Checking Tools Suggested by Tullis and Albert

For Calero et al. accessibility is the extent to which the portal provides enough navigation mechanisms for visitors to reach their desired data faster and easier.⁸ Other researchers also add the language accessibility (use of languages other than the country's own national language) as an important dimension for datasets to be more accessible.⁴¹ Bogdanović-Dinić et al. evaluate the accessibility of open data portals through license and the downloadable feature, i.e. (i) if a dataset is published under an open license, it is accessible to everyone equally, and (ii) if it is downloadable without additional conditions, it is also equally accessible to anyone.⁴⁵

Metric for “Be Accessible”

To assess the accessibility of portals, we developed a metric based on some guidelines and previous studies mentioned in this report. Contrary to previous metrics, in this one we measure each item with a “never, sometimes, always” scale:




- The portal uses human and machine-readable and non-proprietary formats (e.g. CSV, XML, RDF-based formats).
- The portal provides different types of formats for the same dataset.
- The mechanisms for accessing and interacting with datasets are documented.
- Multilingual support is available on the portal.
- The portal supports the visually and hearing impaired.

4 Assessment

After defining our metrics, we selected 10 open data portals to assess their conformity to the 10 user-oriented sustainability principles. By applying the metrics we were able to assess whether they were effective and useful. We considered EU government data portals from different stages of Open Data Maturity (trend-setters, fast-trackers, followers, beginners)¹², as well as some other open data portals from across the spectrum of portal activity.

<p>Cyprus National Data Portal (trend-setter) ¹³</p> 	<p>Avoindata.fi (fast-tracker) ¹⁴</p> 
<p>Data.gov Belgium (fast-tracker) ¹⁵</p> 	<p>Data.gov Slovakia (fast-tracker) ¹⁶</p> 
<p>Dados.gov Portugal (follower) ¹⁷</p> 	<p>Island.is (beginner) ¹⁸</p> 

+

<p>EU Open Data Portal ¹⁹</p> 	<p>London Datastore ²⁰</p> <p>LONDON DATASTORE</p>
<p>Geo Data Portal Luxembourg ²¹</p> 	<p>Open Data Trento ²²</p> 

Next, we discuss in detail the principles and metrics used.

4.1 Organise for Use

The application of this metric was intuitive, and we did not have many problems in using it for the assessment (Table 2). There was quite a lot of variation in the final scores, but no portal achieved the maximum rating. All portals have datasets accompanied by descriptive records and most of them allow to preview extracts of the datasets. Regarding recommendations for related datasets, only one portal (London Datastore) provides this feature. Finland and Slovakia are the only portals that allow users to rate in some way the available datasets. Finally, almost all portals (excepting one - Belgium) provide keywords from datasets that are linked to other published datasets.

Table 2 Metrics Assessment: Organise for Use

Open Data Portal	Organise for Use (out of 5)
EU Open Data Portal	3
Dados.gov Portugal	3
London Datastore	4
Cyprus National Data Portal	3
Open Data Trento	3
Geo Data Portal Luxembourg	2
Data.gov Belgium	1
Avoindata.fi	3
Data.gov Slovakia	4
Island.is	3

4.2 Promote Use

For the principle of “Promote Use”, the metric was also clear enough and easy to apply. In this case, many portals achieved the maximum classification for this metric (**Error! Reference source not found.**). Only one portal had no clear connection with social networks (Belgium). Most of the portals provide users with some kind of online support (e.g. for feedback, to suggest new datasets), though at the time of assessment we were not able to find any relevant support for the case of Iceland (they provide only a contact email and telephone number but there are no FAQs, contact forms, etc.). The majority of the portals also provides a way for users to keep informed of updates to the data, by using RSS feeds, news feeds, and activity streams of the datasets, or by allowing users to “follow” specific datasets of interest. All portals provide links or resources for user guidance and support, by allowing users to contact data producers and/or connect with other re-users. Finally, only 3 portals do not provide examples of reuse of the datasets.

Table 3 Metrics Assessment - Promote Use

Open Data Portal	Promote Use (out of 5)
EU Open Data Portal	4
Dados.gov Portugal	5
London Datastore	4
Cyprus National Data Portal	4
Open Data Trento	5
Geo Data Portal Luxembourg	5
Data.gov Belgium	3
Avoindata.fi	5
Data.gov Slovakia	5
Island.is	3

4.3 Be Discoverable

In this case, all portals achieved the same score (Table 4) i.e. the publisher/owner of the data has an open data portal and publishes an updated, searchable list of datasets. However, it was hard to find if the portals use synonyms while searching for datasets. In addition, we were not able to find any of the selected portals that had the feature of publishing datasets which are known to exist but are not currently available/released. A good example of such feature can be seen at the UK's open data portal (which was not reviewed in this assessment). While a similar score for all portals may mean that the metric is not sufficiently discriminative, in this case, we believe it is showing accurately that portals are not performing strongly in this area.

Table 4 Metrics Assessment - Be Discoverable

Open Data Portal	Be Discoverable (out of 4)
EU Open Data Portal	2
Dados.gov Portugal	2
London Datastore	2
Cyprus National Data Portal	2
Open Data Trento	2
Geo Data Portal Luxembourg	2
Data.gov Belgium	2
Avoindata.fi	2
Data.gov Slovakia	2
Island.is	2

4.4 Publish Metadata

Regarding the assessment of the 5-level maturity schema for publishing metadata (Table 5), 2 portals (EU Open Data Portal and Cyprus) achieved the 5-star classification of Linked Open Metadata, meaning that these portals are applying linked metadata policies, including use of RDF to document their

metadata and also persistent design, use and maintenance of URIs, linking to external vocabularies. Some other assessed portals still are at the stage of having metadata closed or scattered.

Table 5 Metrics Assessment - Publish Metadata

Open Data Portal	Publish Metadata (*)
EU Open Data Portal	★★★★★
Dados.gov Portugal	★★
London Datastore	★★
Cyprus National Data Portal	★★★★★
Open Data Trento	★★★★
Geo Data Portal Luxembourg	★★
Data.gov Belgium	★★
Avoindata.fi	★★★★★
Data.gov Slovakia	★★★★★
Island.is	★★★★★

4.5 Promote Standards

In the case of this metric we found extensive variation in terms of ratings, with some portals having classifications of “4” and others of “1” (Table 6) The metric used for this principle is more difficult to apply because it requires the person who is doing the assessment to have knowledge on RDF vocabulary, such as DCAT, and data management systems for open data, such as CKAN and DKAN. Regarding the use of datasets versioning, we were only able to find one case that uses this feature (Slovakia). However, this is understandable, as having different versions of the datasets would require extra effort and additional operating costs for portal owners. According to Opquast,⁹ viewing these different versions allows the observation of dynamics. However, they also mention that *“the implementation of such a service can quickly introduce significant operating costs. In the case of data that is very frequently updated (or in real time), this feature does not make sense. However, a snapshot can be performed at regular intervals and made available as a version of this dataset.”* Finally, to consider dates that are in a standard format we considered dates expressed according to ISO 8601 (e.g. 2019-02-27).⁴⁶

Table 6 Metrics Assessment - Promote Standards

Open Data Portal	Promote Standards (out of 5)
EU Open Data Portal	4
Dados.gov Portugal	1
London Datastore	1
Cyprus National Data Portal	4
Open Data Trento	3
Geo Data Portal Luxembourg	1
Data.gov Belgium	3
Avoindata.fi	3
Data.gov Slovakia	4
Island.is	3

4.6 Co-locate Documentation

The application of this metric was not straightforward, as it was difficult to understand if the supporting documentation can be immediately accessed from within the dataset. Here, we found classifications of “2” and “3” (Table 7) meaning that all portals provide some kind of supporting documentation but either as a document which has to be found separately from the data, or this documentation is found at the same time as the data but not immediately accessed from within the dataset.

Table 7 Metrics Assessment - Co-locate Documentation

Open Data Portal	Co-locate Documentation (out of 5)
EU Open Data Portal	3
Dados.gov Portugal	2
London Datastore	2
Cyprus National Data Portal	2
Open Data Trento	3
Geo Data Portal Luxembourg	3
Data.gov Belgium	2
Avoindata.fi	2
Data.gov Slovakia	3
Island.is	2

4.7 Link Data

By applying the 5-stars scheme for Linked Open Data, we discovered that while some portals already are positioned at the 4-star level of using RDF standards, none that we reviewed were at the 5-star level (Table 8). The remainder achieved the 3-star rating, as non-proprietary formats (e.g. CSV) were being used.

Table 8 Metrics Assessment - Link Data

Open Data Portal	Link Data (*)
EU Open Data Portal	★★★★★
Dados.gov Portugal	★★★
London Datastore	★★★
Cyprus National Data Portal	★★★★★
Open Data Trento	★★★
Geo Data Portal Luxembourg	★★★
Data.gov Belgium	★★★
Avoindata.fi	★★★★★
Data.gov Slovakia	★★★★★
Island.is	★★★

4.8 Be Measurable

The metric used for the principle of “Be Measurable” was more or less easy to apply and we registered some differences in terms of scores on the portals assessed (Table 9). In some portals we were not able to find relevant analytics, however, this may be because these analytics are not made available for the public. In other cases, we found Site Analytics and Use Analytics. We were able to assess Use Analytics by looking at the cases of reuses of the datasets. However, no portals published Impact Analytics and it was not clear how to assess them independently, therefore this is something that we need to consider for further improvements of this metric.

Table 9 Metrics Assessment - Be Measurable

Open Data Portal	Be Measurable (out of 4)
EU Open Data Portal	2
Dados.gov Portugal	3
London Datastore	1
Cyprus National Data Portal	2
Open Data Trento	1
Geo Data Portal Luxembourg	3
Data.gov Belgium	1
Avoindata.fi	2
Data.gov Slovakia	1
Island.is	1

4.9 Co-locate Tools

By using the metric developed for this principle of “Co-locate Tools”, we were able to find different types of visualisation tools (e.g. maps, graphs, and tables) and collaboration tools (e.g. users posting their reuses, discussion groups and community resources) that are being used by the selected open data portals (Table 10). In this case, Portugal and Luxembourg attained the maximum classification, meaning that the respective portals provided visualisation and collaborations tools to enable users to collaborate innovatively with other users. On the other hand, for Belgium and Iceland we were not able to find relevant visualisation or collaboration tools.

Table 10 Metrics Assessment - Co-locate Tools

Open Data Portal	Co-locate Tools (out of 4)
EU Open Data Portal	2
Dados.gov Portugal	4
London Datastore	3
Cyprus National Data Portal	2
Open Data Trento	2
Geo Data Portal Luxembourg	4
Data.gov Belgium	1
Avoindata.fi	3
Data.gov Slovakia	3
Island.is	1

4.10 Be Accessible

For the case of “Be Accessible”, in general the ratings were good (**Error! Reference source not found.**). By applying our metric, we felt the need to add a new item for the scale of “never, sometimes, always”, by adding “most of the time”. This change allows to differentiate further between portals. All portals use human and machine-readable and non-proprietary formats, although some are “not always” or just “sometimes” doing it. The same happens for the provision of different types of formats for the same dataset. In addition, all portals are in some way documenting mechanisms for accessing and interacting with datasets. Iceland and London Datastore are the only portals providing support in only one language (Icelandic and English, respectively). Finally, we were not able to find portals that support the visually and hearing impaired.

Table 11 Metrics Assessment - Be Accessible

Open Data Portal	Be Accessible
EU Open Data Portal	Most of the time, Most of the time, Always, Always, Never
Dados.gov Portugal	Sometimes, Most of the time, Always, Always, Never
London Datastore	Sometimes, Sometimes, Always, Never, Never
Cyprus National Data Portal	Most of the time, Sometimes, Always, Always, Never
Open Data Trento	Most of the time, Most of the time, Always, Always, Never
Geo Data Portal Luxembourg	Sometimes, Most of the time, Always, Always, Never
Data.gov Belgium	Sometimes, Sometimes, Most of the time, Always, Never
Avoindata.fi	Sometimes, Sometimes, Always, Always, Never
Data.gov Slovakia	Most of the time, Most of the time, Always, Always, Never
Island.is	Sometimes, Sometimes, Always, Never, Never

5 Recommendations

Portal owners should consider a regular assessment of their portals along these dimensions using the metrics above, as part of a strategic assessment of what they would like to achieve.

Areas with particularly low diversity of results, such as Be Discoverable and Co-locate Documentation, should particularly be addressed. Are there technical or social barriers preventing the implementation of improved solutions?

Recent research has looked at ways to automatically assess some of these metrics and analysed a subset of (CKAN-based) portals indexed by the EDP.² Among other things, the analysis showed that current technical realisations of portals do not lend themselves well to a continuous, detailed monitoring of data use, which in turn means that portal owners have limited insight into the impact of their publishing effort. Further work should be invested in this area.

6 Conclusion

Amongst the variety of open data assessment tools and methods discussed in the introduction, it is clear there is a gap regarding the usability of open data portals that requires addressing, and this report sets out to fill that gap.

It does so firstly by reviewing existing research, and developing metrics for assessing how sustainable the portals are, according to the Analytical Report 8, Future of Open Data Portals. Secondly, it applies those metrics to ensure that they can be used successfully in real portal scenarios.

Regarding the application of the metrics, we conclude that not all of them are equal in effort to apply. Naturally, some are more challenging than others, requiring more time to do the portal assessment, and sometimes requiring a previously acquired knowledge on this area of Open Data, such as RDF vocabulary and data management systems for open data.

Nevertheless, these issues should be easily overcome by the portal owners. Together, these user-oriented sustainability principles and metrics represent a coherent strategy to minimise current problems and achieve use and impact.

Results from this assessment showed that portals are doing well in some areas but should reflect on their strategy in other specific areas. Naturally, some portals are already well developed with high

² Dix, 2019

levels of maturity for open data, while others are still in a process of development. In general, we found that the more mature portals were already applying many of the higher level principles.

Portal owners can use these metrics not only to assess the current sustainability of their portal, but also to identify exactly what steps they need to take to improve their usability. They may decide to focus on areas where they obtained the lowest scores, or on areas where other portals score more highly, and offer examples that can be replicated.

7 Referenced Studies

¹ Walker and Simperl (2017):

https://www.europeandataportal.eu/sites/default/files/edp_analyticalreport_n8.pdf

² Umbrich et al (2015): <https://doi.org/10.1109/FiCloud.2015.82>

³ European Data Portal (2015):

https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n1_-_final.pdf

⁴ <https://index.okfn.org>

⁵ <https://census.okfn.org/>

⁶ <https://opendatabarometer.org/>

⁷ <https://theodi.org/article/open-data-maturity-model-2/>

⁸ Calero et al (2008): <https://doi.org/10.1007/s11280-008-0048-y>

⁹ <http://checklists.opquast.com/en/opendata/>

¹⁰ Welle Donker and van Loenen (2017): <https://doi.org/10.1080/17538947.2016.1224938>

¹¹ Colpaert et al. (2013): <http://pieter.pm/5stardataportals/>

¹² European Data Portal (2018):

https://www.europeandataportal.eu/sites/default/files/edp_landscaping_insight_report_n4_2018.pdf

¹³ <https://data.gov.cy/?language=en>

¹⁴ <https://www.avoindata.fi/en>

¹⁵ <https://data.gov.be/en>

¹⁶ <https://data.gov.sk/en/>

¹⁷ <https://dados.gov.pt/en/>

¹⁸ <https://opingogn.is>

¹⁹ <https://data.europa.eu/euodp/en/home>

²⁰ <https://data.london.gov.uk>

²¹ <https://www.geoportail.lu/en/>

²² https://dati.trentino.it/en_GB/

²³ Tullis and Albert (2013): <https://www.elsevier.com/books/measuring-the-user-experience/albert/978-0-12-415781-1>

²⁴ <https://www.measuringux.com>

²⁵ <http://opquast.com/en/>

²⁶ Lin & Yang (2014): https://doi.org/10.1007/978-3-662-45071-0_26

- ²⁷ Alexopoulos et al. (2014): https://doi.org/10.1007/978-3-662-44426-9_19
- ²⁸ Máchová & Lněnička (2017): <https://doi.org/10.4067/S0718-18762017000100003>
- ²⁹ Goh et al. (2008): <https://doi.org/10.1108/14684520810889664>
- ³⁰ Attard et al (2015): <https://doi.org/10.1016/j.giq.2015.07.006>
- ³¹ Walker et al. (2015):
<https://eprints.soton.ac.uk/375700/1/user%2520centred%2520metrics%2520final.pdf>
- ³² Maali et al (2010): https://doi.org/10.1007/978-3-642-14799-9_29
- ³³ Máchová et al (2018): <https://doi.org/10.1108/AJIM-02-2018-0026>
- ³⁴ <https://5stardata.info/en/>
- ³⁵ European Commission (2011): https://joinup.ec.europa.eu/sites/default/files/document/2011-12/towards_open_government_metadata_0.pdf
- ³⁶ <https://joinup.ec.europa.eu/solution/dcat-application-profile-data-portals-europe>
- ³⁷ <https://www.w3.org/TR/gov-data/>
- ³⁸ Sulova (2014): <https://ideas.repec.org/a/vrn/journal/y2014i3p37-47.html>
- ³⁹ <https://data.europa.eu/euodp/en/visualisation-home>
- ⁴⁰ <https://ec.europa.eu/eurostat/help/first-visit/tools>
- ⁴¹ Sayogo et al (2014): <https://doi.org/10.1109/HICSS.2014.240>
- ⁴² <https://opengovdata.org>
- ⁴³ <https://opengovdata.io>
- ⁴⁴ <https://www.w3.org/TR/WCAG20/>
- ⁴⁵ Bogdanović-Dinić et al (2014): https://doi.org/10.1007/978-1-4614-9982-4_3
- ⁴⁶ <https://www.iso.org/iso-8601-date-and-time-format.html>