Sustainability of (Open) Data Portal Infrastructures

Measuring Use and Impact of Portals

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:
Mark Frank
Email: mark.t.frank@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

OA-02-20-164-EN-N ISBN: 978-92-78-42147-2 doi: 10.2830/851783

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

Table of Contents

1. Introduction

1.1 Scope of this report

Measuring the value of open data is a broad subject that has generated a lot of research and discussion. This paper is focussed on three key aspects:

- **Measuring the *economic* impact of open data portals,** as opposed to the cultural, social or political value;
- **Measuring the impact of open data *portals* (including the data they contain).** It does not focus on measuring the value of open data in a country, such as the Open Data Index or Barometer seek to do;
- **Measuring the *impact* of open data portals.** It follows the CAF framework in separating use and impact, and examines use in order to signpost impact.

This report forms part of a larger report on Sustainability of (Open) Data Portals. It aims to locate and disambiguate what is being measured, and understand which of the methods currently and potentially available to us, are the most appropriate to utilise. It also provides a basis for the report T1.2: Developing Microeconomic Indicators Through Open Data Re-use.

1.2 Aims of this report

This paper builds on previous studies into the economic impact of open data portals, in particular the Open Data Portal Recommendations – From Set up to Sustainability[1] (ODPRSS 2017) and the subsequent report Ensuring the Economic Sustainability of Open Data Portals[2] (EESODP 2018). The latter report explores the role of the measurement of economic impact of an open data portal in creating a sustainable environment for the portal, and takes an overview of possible methods. After reviewing these two reports, the paper:

- Critically examines in more detail which attributes of open data portals we should be measuring and the process for choosing what to measure;
- Explores which methods can be used to measure these more specific attributes;
- Assesses these methods using criteria proposed by Frank and Walker.[3]

Exploring the attributes to be measured in more detail opens up a more detailed analysis of current and potential future measurement methods, which creates the requirement for a systematic development and assessment of those methods. This is the focus of further work in this Task., which links the questions about open data - who is using it? What impact is it having? - to answers in the form of indicators and metrics. The intended result is an evidence-based model that describes the relationship between source and use of data, and identifies the bases of economic impact.

[1] Recommendations for Open Data Portals: from setup to sustainability (2017)– chapter 7

[2] Ensuring the Economic Sustainability of Open Data Portals (2018)– chapter 2

[3] Frank M. and Walker J. (2015). User Centred Methods for Measuring the Value of Open Data. *The Journal of Community Informatics*, *12*(2).

1.3 Terminology

Throughout this paper, three terms are used repeatedly with specific meanings:

Attribute refers to a property of interest. For example, the portal management may be interested to know what proportion of visitors to a portal work in particular sector as this may affect the choice of data to be published.

Metric refers to a property that is measured. For example, the portal management may measure percentage of visitors to the portal who report they work in that sector.

Method refers to the process for measuring one or more metrics. For example, a self-selecting survey of users.

1.4 Structure of this Report

We firstly set the context of previous work, including building on the recommendation to 'adopt and adapt the Common Assessment Framework' for measurement. This report then demonstrates how we can use this to define attributes (things we want to measure). We then review the methods available for assessing these attributes. In doing so, we demonstrate that the choice of metric or indicator for each attribute is largely dictated by the method used to measure the attribute. We then assess which method is likely to be the most appropriate for measuring each attribute, in terms of producing the most useful, reliable, transferable, comparable and valid metrics.

1.5 Methodology

Based on the previous work, we researched the Common Assessment Framework further;

We updated the methods where appropriate with new research since the publication of the previous reports;

We researched existing metrics that could be applied from other, long-standing statistical and data sources;

We presented the work on the relevant sections of the Common Assessment Framework and associated attributes, methods for measurement and metrics to a workshop of National Portal Representatives for their input and feedback.

2. Relation to previous work

The report ODPRSS focuses on methods that can be relatively easily automated using currently available technology, such as:

- **What data has been published**. For example, a simple count of the number of datasets that have been published on a portal, sometimes classified by subject matter, or by technical criteria such as format;
- **The quality of the data.** For example, the proportion or number of datasets that are published with acceptable open licensing arrangements, or with an acceptable completeness of metadata. The latter typically reports on which metadata fields have been filled in and does not attempt assess the quality of the metadata e.g. whether it is true or even makes sense;

- **The number of visitors to the portal and information about them that can be obtained from portal logs.** For example, which browser or device are they using to access the portal and in which country are they located;
- **Access to data.** For example, which pages on the portal do visitors look at overall and the number of downloads for specific datasets.

These measurements are very useful. They help assess:

1. The rigour and efficiency of the publishing process, e.g. data quality reports will indicate if the process is publishing data with an appropriate license.
2. The design of the portal, e.g. an analysis of pages accessed may show how easily users are able to find datasets.
3. The choice of datasets, e.g. portal logs may show if there are many datasets that are rarely accessed.

However, they are of limited value in assessing the economic impact of the portal, or of the data available through that portal. The subsequent EESODP report surveys a much broader range of measurement methods that possess the potential to measure economic impact, introduces some important conceptual distinctions, and makes some significant recommendations for portals. This paper builds on this, by critically analysing the practical implications of those recommendations.

The EESODP report is a high level analysis. It bases its discussion of the attributes to be measured on the Common Access Framework[4] (CAF) – a joint project between several global institutions seeking to standardise the measurement of open data. It makes the recommendation to "Adopt and adapt the Common Assessment Framework to measure portal performance, identifying and using relevant existing metrics around Data and Context/Environment."[5]. This report analyses the attributes in CAF then provides a critique of these to develop a coherent overview of attributes of interest. The paper then presents a detailed analysis of present and future measurement methods for metrics associated with these attributes. It then goes on to analyse the strengths and weaknesses of each method as applied to different attributes using a systematic approach based on criteria proposed by Frank and Walker (2016).

[4]http://opendataresearch.org/sites/default/files/posts/Common%20Assessment%20Workshop%20Report.pdf
[5] Ensuring the Economic Sustainability of Open Data Portals - recommendations

3. The Common Assessment Framework (CAF)

3.1 High Level Categories

Context	Data	Use	Impact
Legal	Licensing – how open Technical – format, APIs, documentation What data – core data, sectors represented Quality – up to date, complete	Type of users – researchers, entrepreneurs Purpose – reduce spending, ease congestion Activities – benchmarking, mapping	Environmental - reduced pollution
Organisational			Economic - increased jobs, growth
Political			Political - reduced corruption, better services
Legal			Social - greater equality, participation
Social			
Economic			

Figure 1 High Level Overview of the Common Assessment Framework

The EESODP report uses slightly different high level categories to CAF. While CAF classifies attributes under Context, Data, Use or Impact, the report sees Context as largely irrelevant to portals because there is little need to compare the performance of one portal against another in a different context. It renames Data as Access to place the emphasis on measuring the success of the activity of accessing the data – rather than simply measuring the data. As a result, EESODP recognises three high-level categories:

Access – the act of retrieving data in machine readable format.

Use – the act of processing the data in order to create information from which decisions can be made.

Impact – the outcome of decisions made based on information derived from open data.

However, CAF includes further detail. There is a report on the initial CAF workshop that breaks down the range of attributes to some extent[6]. Subsequently this has been informally updated through a

[6] Towards common methods for assessing open data: workshop report & draft framework. http://opendataresearch.org/sites/default/files/posts/Common%20Assessment%20Workshop%20Report.pdf

Google document [7], most significantly by providing examples of measurement in practice. The most recent update at the time of writing was July 2018. Below, we provide an overview of the attributes of each category.

3.1.1 Data/Access

The CAF data attribute (which corresponds to the EESODP access attribute) is broken down into three core questions each of which suggests more specific attributes which we may want to measure.

***Dimensions**: What are the technical, legal, practical, and social dimensions of openness?*
These are common aspects of data, which are independent of the content and quality. They include:

- technical dimensions such as the format and structure of the data, the availability of APIs and the quality of metadata;
- social dimensions such as quality of documentation and existence of support;
- legal dimensions such as availability of appropriate licenses;
- practical dimensions such as how easily is the data discovered.

***Sector/classification of datasets**: What kinds of datasets are available within a country or sector?*
These attributes are based on the content of the datasets. For example, which industry sectors are covered, and the availability of core datasets.

***Quality:** How complete, primary, timely, usable and reliable are specific datasets, or the open data provided in general?*
These attributes primarily reflect the processes used to publish and maintain the data, although usability, for example, may also be a function of the portal design.

The Data/Access attributes are therefore addressed through current automated methods (more detail on which can be found in ODPRSS 2017). However, there is a potential relationship between these attributes and their use which we propose to investigate further.

3.1.2 Use

CAF breaks down the attribute of use into three questions.

***Users:** Who is using open data?*

This is typically interpreted as a job or role e.g. researchers, entrepreneurs although other attributes may be relevant such as demographic information (age, class, ethnicity).

***Purposes:** For what purposes are individuals/organisations using open data?*

Typically expressed in terms of business objectives such as reducing spending or increasing customer satisfaction.

***Activities:** To what uses are individuals/organisations putting open data?*

How is the data actually being used e.g. benchmarking and commercial applications.

[7] https://docs.google.com/document/d/1DLQrC-UnvK_3-aVGMB0AS1zqNHbZ6NxUcAfqY8ksETc/edit

3.1.3 Impact

CAF breaks down the impact into Social, Environmental, Political and Economic/commercial**.** The four key questions are:

Social: *What are the social benefits to be gained from the use of open data?*

Environment: *What are the potential benefits of open data for environmental sustainability?*

Political/Governance: *How does open data help improve government efficiency and accountability?*

Economic/Commercial*: What are the impacts of open data on economic growth and innovation?*

<table>
<tr><th>Data/Access</th><th>Use</th><th>Impact</th></tr>
<tr><td rowspan="6">Dimensions – legal, technical, practical and social

Sectors/classification - core data, sectors represented

Quality – up to date, complete</td><td rowspan="6">Type of users – researchers, entrepreneurs

Purpose – reduce spending, ease congestion

Activities – benchmarking, mapping</td><td>Environmental - reduced pollution</td></tr>
<tr><td>Economic - increased jobs, growth</td></tr>
<tr><td>Political - reduced corruption, better services</td></tr>
<tr><td>Social - greater equality, participation</td></tr>
<tr><td></td></tr>
<tr><td></td></tr>
</table>

Figure 2 CAF as Configured for Portals (EESODP 2018)

This report is focussed on economic/commercial impact. The CAF examples concentrate on the macroeconomic impact e.g. economic growth. However, it is also relevant to consider the microeconomic impact. The microeconomic impact is closely tied to the purpose, but with the emphasis on measuring to what extent the purpose has been fulfilled.

3.2 Critiques of the CAF

When using the CAF to identify possible indicators and metrics it is important to understand what the alternatives might be. Although CAF creates a framework for measuring different dimensions of open data, it is focussed on a single dimension at a time. A different approach is to assess which business

models are supported by the portal and the data therein. A business model has multiple dimensions. For example, Lindman et al[8] propose five open data business models based on four dimensions:

- What service is being offered – this is similar to user activity in the CAF framework
- The revenue model – how is the service funded
- Resources – what resources are being used to perform the service
- Relationships – how does the service use external resources

Two further aspects of CAF complicate its applicability. The first of these is its macro-level approach. It understands measuring impact in terms of economy-level growth and innovation, much as is addressed by national-scale studies such as the Open Data Barometer and Open Data Index, and less so in terms of regional or specialist portals which are looking for impact on a smaller, more specific scale.

CAF is undoubtedly useful in separating use and impact in terms of demonstrating these are two different things and that use does not directly imply impact. This does not mean that we cannot derive impact metrics from use attributes, as there is unlikely to be impact without use. An example is the 'activities' attribute 'benchmarking'. If there was a substantial uptake of open data for benchmarking it would make sense to try to identify the impact of this. There is also a relationship between access and impact: if data on specific themes is frequently accessed it is logical to look for impact in related areas (this approach is enlarged upon in Task 2: Automated Assessment of Indicators and Metrics).

4. Selecting Attributes to Measure

When selecting attributes, CAF has limitations. It is limited to the four top-level categories of attributes (three considered in this report) and the subcategories immediately below. Furthermore, there is no explicit basis for choosing these subcategories or for expanding them, other than the experience of the participants. Fortunately, there are precedents outside of open data, which can act as a guide to selecting more specific attributes of interest. If open data is to have an economic impact, it does so by providing information to users. There is a long history of measuring the impact of other sources of information, such as weather forecasts and census data. Such sources of information can indicate useful and appropriate attributes and some methods for measuring them. Using established attributes like these has several advantages.

- Their significance, strengths, and weaknesses are well understood;
- The link between information, these attributes, and the associated measurement methods has been explored extensively;
- They facilitate comparison between open data and other information source
- In many cases, there are international standards that facilitate comparison across countries.

A weakness of this approach is that it may be limited to measuring the economic impact of one type of information. The appropriate attributes for assessing the economic impact of information about the weather are unlikely to be the same as those for measuring the impact of census information

[8] Lindman et al, 2014, Industrial Open Data: Case studies of early open data entrepreneurs, 47th Hawaii International Conference on System Science

about household composition. The data on a typical portal is of many different types. So the choice of attributes for a particular source of information may be appropriate for assessing the impact of some specific data on a portal but not for assessing the portal as a whole. An exception is attributes and metrics that fall under the CAF data\access heading. These are unique to open data and do not apply to other information sources.

It is important to choose attributes that it is possible to measure i.e. to choose attributes that have methods that are, at a minimum, valid and reliable. This implies that the choice of attributes and metrics cannot be isolated from each other. They must be considered together.

5. Methods

The EESODP report identifies five generic methods for measuring open data:

5.1 Macroeconomic Studies

These use established econometric methods and are concerned with "the performance, structure and behaviour of an economy as a whole" [9]. Because they use established methods, interested parties such as government, business leaders and civil understand their significance and the results are credible. However, they are expensive to run and for the most part are intended to assess in general terms the impact of open data on a national economy over several years.

5.2 Microeconomic Studies

These also use established econometric methods but focus on specific publishers or datasets. They share the same strengths and weaknesses as macroeconomic studies: the results are credible and stakeholders understand their significance but they are expensive and hard to repeat. However, one reason this is true is because there is no commonly agreed set of indicators that are being measured, and so each study acts as a standalone. Creating a transferable set of metrics and indicators that could be reused would assist with this (for more on this, see the report 'Creating Microeconomic Indicators Through Open Data Re-use'). One such example is the report 'Assessing the Value of TfL's Open Data and Digital Partnership'. This identified direct benefits, realised in the form of revenues from market transactions and indirect benefits of positive externalities, for example, increased engagement with municipality and services. These are reliable and comparative metrics.

While expensive, some of the cost of running such surveys can be defrayed by incorporating portal surveys into larger studies run by business associations, such as that run by ASEDIE, the Spanish Multisectorial Information Association. They conduct longitudinal analyses of the information and data market, and the outputs of trade and governmental economic assessments can be useful to local assessments of the impact of open data in a local marketplace.

5.3 Business Population Studies and User Surveys

Unlike macroeconomic studies and microeconomic studies, these are aimed at the users of the open data from a portal and are intended to understand how open data is being used and thus assess its

[9] https://www.europeandataportal.eu/sites/default/files/s3wp4_sustainability_recommendations_ii.pdf p32

impact. They “tend to focus on making a holistic assessment of the overall ecosystem” revealing “what types of data are most used and how that data is being used”. They typically gather information using surveys and publicly available sources of data such as company registers. They provide an important different perspective from economic studies but also suffer from being expensive to run and therefore hard to repeat. Additionally, because they rely so much on self-reporting surveys, there are often concerns over whether they have obtained a reasonable representation of all users of the data. Again, a common framework of metrics and indicators would assist with this.

5.4 Showcases and Use Cases

These are examples of the use of open data which are often displayed on portals. They typically rely heavily on self-reporting by users. They are a valuable source of information about the use and impact of the open data on a portal and can be less expensive than the previous three methods (the cost depends on the extent to which the portal management decide to manage and market the recording of use cases). They tend to get out-of-date, sometimes lack credibility, and by definition only address a few examples of the use of data.

Despite this, there are now a number of substantial corpuses of use cases and showcases. There are over 550 use cases on the European Data Portal, 1793 on data.gouv.fr and 232 on datos.gob.es. ODImpact.org is a site devoted to such use cases. While these were originally intended to understand how impact might be derived, the range of uses, and of course, to inspire, they are increasingly available in numbers that can be analysed quantitatively to create metrics. These might include:

- Number and quantity of data themes
- Number of types of reuses
- Log files
- Inferring user needs from quantifying areas of interest –If they need it, it should be impactful
- Measuring impact of hacks from apps developed
- Speed of addition to portal (as a proxy for rates of reuse)

Exploring individual applications at a more micro level can also help to measure impact. For example, a Journey Planner App for bikes might be able to demonstrate impact via the size of the installed and user bases. Obviously, these indicators have to be obtained, which can be facilitated by publishers and reusers working together. One suggestion to enable this was requesting that showcased app developers commit to report on a set of indicators in exchange for promotion of the app.

5.5 Automated Access Metrics

These use portal technology to automatically measure various aspects of visitors accessing the data on the portal. They include:

Page analytics. These are similar to other types of web site, which record metrics such as which pages are accessed most often and the order in which they accessed.

Downloads. Which datasets are downloaded, and how often.

API metrics. Where the portal enables users to access data through an API it is possible to record how often the API is used and some data about who is using it.

Because they are automated, these metrics are less expensive than the other four types of method and can be implemented on an on-going basis. Although there can be technical limitations, depending on portal software, there are also many established tools for analysing and reporting on this type of data. There is also considerable scope for common practice amongst portals which potentially allows sharing and comparison of data. The biggest limitation is that they are restricted to measuring how the data is accessed and provide little information on its use and impact. They can throw light on data attributes but not the others.

5.6 Additional Methods

This report proposes a modified version of this list of five methods.

5.6.1 User Surveys

It separates **user surveys** from **business population studies**. Business population studies assess the "overall ecosystem" of the users of the portal data (such studies may be include one-off customised user surveys). This is an important method, but it is relatively expensive as it means gathering information about the users from multiple sources, possibly creating surveys that are specific to that study, and using the information to create an overall picture of how the data is being used. Often some of the sources of data are specific to that portal e.g. data on organisations in the geography served by the portal that is not available in other geographies. It is essentially a one-off snapshot of the users at the time of the study. There is potential for stand-alone user surveys of all users of the portal, which are done semi-automatically, at low cost, and continuously, e.g by integrating them into portal access or the data itself. This approach is significantly different from business population studies and merits separate analysis.

User surveys can play an important role intermediate between automated methods and more resource intensive methods such as business population studies and microeconomic studies. Unlike current automated methods, they can address user attributes. At the same time, they have the potential to be semi-automated and therefore low resource. For example, a portal may require users to register and thus gain information about them. Alternatively, surveys can be integrated into the data itself. This could be as simple as an additional page in a spreadsheet or a link to a survey embedded in the data.

These are aimed at the users of the open data around a portal and are intended to understand how open data is being used and thus assess its impact. They typically gather information using surveys and publicly available sources of data such as company registers. They provide an important different perspective from economic studies but can also suffer from being expensive to run and therefore hard to repeat. In addition, because they rely so much on self-reporting surveys, there are often concerns over whether they have obtained a reasonable representation of all users of the data. However, where they are used, they are some of the most impactful and useful studies.

A (comparatively) simple example is offered by the Irish national portal, which runs a continuous on-portal user survey offered to all users who download a dataset, which allows both engagement and measurement. In the future, new automated methods and social media analysis may avoid the need for compromise in some contexts, but this is currently some way off

Although user surveys have significant potential, there are also drawbacks.

- They may be considered to be intrusive and contrary to the spirit of open data;
- They are likely to be limited to the primary users of the data as they are the only ones likely to be confronted with the survey. It may well be that the most significant economic value lies with secondary or tertiary users;
- To gain the advantages of automation, users need to complete the surveys without human assistance. This means that they are likely to be self-selecting;
- It is vital that the survey questions are clear and unambiguous. This is easier in some contexts than others. For example, some industries have clearer job roles than others.

Nevertheless – user surveys are sufficiently distinctive and potentially useful to merit separate consideration.

This paper also includes two potential, as yet untried methods. These are discussed in the EESODP paper but are not analysed alongside the current methods.

5.6.2 Future Automated Methods

Future automated methods refers to automated methods that are the subject of discussion and research but are not yet in use. EESODP discusses some leading contenders: tracking users through API keys, tracking use through version control, identifying use through web search, and measuring downstream usage through data citation.

Potentially, such methods can measure a much wider range of attributes than current automated methods. Current automated methods have low resource requirements but are poor at measuring use and impact. Future automated methods have the potential to measure use, and possibly impact, while retaining the low resource requirement. However, future automated methods present both technical and cultural challenges. For example, data citation technologies require users to cite data sources in a consistent way. Technology can make this as straightforward as possible but it still requires a change in practice. Therefore, these methods merit research into both the technology and the cultural aspects of using them. An example of this can be seen in the report, 'Automated Assessment of Indicators and Metrics'.

Social media analysis is another potential but untried approach, which takes advantage of the wide range of existing methods, such as sentiment analysis, used in other fields for automatically analysing social media. This loosely defined set of methods has the potential to measure both use and impact in a meaningful way without using excessive resource. At the time of writing there has been very little experience or research on using social media analysis in the context of open data.

The revised list of methods is listed below. As a consequence of the changes above **automated access metrics** has been renamed **current automated methods** to differentiate it from future automated methods and to make clear it refers to methods not metrics. **Showcases and use cases** have been abbreviated to **case studies**.

The table also summarises the main characteristics of each approach.

Scope. Does this method typically address the users of a portal or users from a particular geography or some other grouping of users?

Data collection. How frequently is data collected using this method? In particular, is it a one-off process, or is there on-going data collection allowing for the measurement of trends and continuous monitoring.

Resource required. How expensive is this method? Expensive methods will usually be one-off.

Approach	Scope	Data collection	Resource required
Current automated methods	Portal	Ongoing	Low
User surveys	Portal	Ongoing (potentially)	Low
Case studies	Portal	Ongoing	Medium
Business population studies	Geographic	One-off	Medium
Microeconomic studies	Publisher/ dataset/ sector	One-off	High
Macroeconomic studies	Geographic/ economic	One-off	High
Future automated methods	Portal	(orOngoing geographic)	Low
Social media analysis	Portal	(orOngoing geographic)	Low

Table 1 Summary of Measurement Methods Reviewed

6. Assessing Methods and Metrics

The choice of method for measuring something is based on several different criteria, depending on what attribute is being measured and the reason for doing the measuring. Frank and Walker have proposed a systematic approach to the assessment of metrics in the context of open data. While the Frank and Walker paper discusses metrics, the same considerations apply to methods. They propose these six criteria for assessing a metric or method:

Method Assessment	Description
Valid	The method is closely correlated with the attribute of interest
Reliable	The method gives consistent results over time and between observers
Sensitive*	The method is sensitive enough to discriminate significant differences in the attribute of interest.
Efficient	The less time and resource required to use it the better. In some contexts, poor efficiency can lead to poor validity and reliability.
Transferable	The same method can be used in a variety of different contexts and across cultural and economic variation.
Comparable	If a method is comparable not only is the method transferable to a wide variety of contexts but the results can be meaningfully compared. Ideally this would result in a universal standard that transcends cultures and applications.

Table 2 Criteria for Assessing Methods and Metrics

* Frank and Walker describe this criterion as discriminatory, which leads to an awkward noun for the criterion: discriminability. This paper has renamed the criterion as sensitivity, so a method which has good sensitivity is sensitive.

Note that some of these criteria depend on others. A method cannot be sensitive unless it is valid and it cannot be comparable unless it is both valid and transferable.

In the context of open data portals, efficiency is of great importance. Measurement of economic impact is of limited value if it is a one-off snapshot. A snapshot can be used to communicate the potential of a portal and justify its existence, but a realistic assessment of its contribution in the longer-term means measuring trends in impact over time. However, it is only feasible to sustain ongoing data collection if the resource required is low. This in turn implies some level of automation.

The next section explores each CAF attribute in more detail and assesses the appropriate current methods for measuring them.

6.1 Data/Access

It is relatively easy to measure most of the data attributes using current automatic methods, or with minimal manual intervention, provided there are comprehensive and well-maintained metadata.

6.1.1 Sector/Classification of Datasets

Sectors can easily be measured automatically if data providers include the sector as part of the metadata, e.g. by simply counting the sectors represented by datasets on the portal. This crude method risks counting datasets that are out of date or not truly representative of the sector. A more sophisticated model might reduce this risk by weighting the count according to how recently the datasets have been updated or accessed.

Key to this is a common understanding of the definitions of the sectors and therefore it is important to use widely accepted definitions. Fortunately, there are commonly used definitions of sectors at

national and European levels independent of open data[10]. The validity, transferability and comparability of the metric will depend on this common understanding and the rigour with which the metadata are supplied and maintained. If this is done well, then such straightforward methods should be sensitive, reliable, and efficient.

The presence of core datasets raises similar considerations. If there are widely accepted definitions of core data, and a suitable process for creating and maintaining metadata, then this can easily be automated. Metrics might simply be the list of core datasets on the portal or a list weighted by how recently they were accessed or updated. There are a number of approaches to identifying such 'core' datasets. The 'High Value Datasets' work currently being undertaken is one such approach. Other lists have been produced by the G8 Open Data Charter, the Open Data Index, the Open Data Barometer and the Open Data 500.[11]

6.1.2 Data Quality

Current automated methods can use metadata to measure some aspects of data quality. A good example is using frequency of updates to measure timeliness[12]. Such methods are efficient and reliable. There are some concerns over validity and therefore sensitivity. For example, it is possible for data to appear to be up-to-date because it is changed frequently, but for the changes to fail to reflect recent events. Also, while these methods can be relatively easily used in other contexts and are therefore transferable, the results are not comparable. For example, an annual update frequency is perfectly acceptable for demographic data on a national portal but quite unacceptable for data on train movements on a transport portal.

Other data quality attributes such as the completeness and usability of the data require more resource intensive methods. It is likely to be necessary for people with subject matter knowledge to inspect the data to determine whether it is complete or has significant omissions. It requires a study or review of the data in use to determine whether it is usable in practice (for example is it the right granularity?). Such studies can be valid, reliable and sensitive, but they are inefficient and hard to transfer to other contexts.

6.2 User

6.2.1 Type of User

This potentially includes a vast range of different attributes. However, as the CAF Google document notes, *"a systematic taxonomy cutting across all sectors is yet to be developed. "* Leading attributes for understanding economic use are:

1 Occupation (researcher, manager, etc)
2 Industry (manufacturing, agriculture, education, government etc)
3 Demographics such as gender, age and socio-economic class.

[10] e.g. https://ec.europa.eu/growth/sectors_en
[11] All these are linked to from the CAF document: https://docs.google.com/document/d/1DLQrC-UnvK_3-aVGMB0AS1zqNHbZ6NxUcAfqY8ksETc/edit
[12] Ulrich Atz (2014) The Tau of Data: A New Metric to Assess the Timeliness of Data in Catalogues. In Proceedings of the International Conference for E-Democracy and Open Government (CeDEM2014), Krems, Austria

There is a great deal of experience of measuring these attributes independently of open data through instruments such as national censuses and social science research. They use well understood, widely used, and, in some cases, standardised, taxonomies and allow the results of any open data metrics to be compared to other data. For example, the International Labour Organisation uses a standard set of occupations. [13] This is just one of many standard international classifications maintained by the United Nations Statistical Division (UNSD).[14]

Several of the methods discussed above can reasonably be used to measure these attributes. Case studies, user surveys and business population studies can all be used to determine the type of user. Microeconomic studies might possibly include determining the type of user although they often work at the level of a business or organisation rather than individual users.

A key approach to facilitate measurement of use is to focus on primary users. Amongst current methods, the user survey is an attractive but hugely underutilised compromise for measuring user type and activity amongst primary users. This has not always been the case, especially with arm's length measurements such as macroeconomic surveys. This can be done by engaging more closely with users. Developing an increased relationship with the community implies a two-way dialogue that will ultimately be beneficial. In this way, the impact can be crowd-sourced in a number of ways

User surveys stand out as being a current method that has high efficiency and therefore the scope for being used on an on-going basis. Validity, reliability and sensitivity are crucially dependent on clear definitions. However, even when standard definitions are used, reliability and validity will be compromised, as user surveys need to be self-selecting and will therefore be incomplete and biased. While such surveys can be transferred to other portals and contexts fairly easily, the comparability of the results will vary depending on the context. For example, the user's choice of industry is likely to be reasonably consistent across a range of geographies and other contexts. The interpretation of "agriculture" is fairly unambiguous. On the other hand, the choice of occupation will rely on using standard definitions and on users understanding those definitions. For example, "Legal, Social and Cultural Professional" may be open to a wide range of interpretations in different contexts.

Case studies, business population studies and microeconomic studies use several sources of information and compare the results. This means they potentially have higher validity, reliability and sensitivity than user surveys. They often take advantage of sources of information that are specific to a portal and therefore have limited transferability and comparability. All these methods use extensive resource and are therefore low on efficiency.

6.2.2 Purpose

As the paper is focussing on economic impact, the relevant attributes are likely to be business objectives such as increase profitability, reduce costs, or decrease staff turnover. This type of attribute is generally well understood, well defined, and relatively consistent across a wide range of sectors, geographies and roles. For the most part the corresponding metric is likely to be a count of the users with particular purposes.

[13] International Standard Classification of Occupations
[14] UNSD Statistical Classifications

The most plausible method for measuring business objectives is again a user survey. User surveys can be efficient and sensitive. However, there are greater concerns about the validity and reliability of user surveys for measuring purpose than for there are for user type. Users are less likely to understand the underlying business purpose of using the data than they are to understand user type. They may simply be responding to directions from management or requests from another department or organisation. The fact that user surveys are usually limited to primary users is of particular concern in the case of purpose. For example, while the purpose of an application provider using transport data may be to increase revenue – the purpose of the users of the application may be to reduce travel time. Therefore, the validity and reliability of user surveys in the context of measuring purpose are medium to low with similar consequences for sensitivity, transferability and comparability.

Case studies, business population studies and microeconomic studies are potentially more valid and reliable than user surveys but these methods have the same problems as they do when measuring user type – they are inefficient and likely to be limited to a subset of the users of the portal.

6.2.3 Activities

Again, user surveys are the most plausible method, out of those listed, for measuring activities on an ongoing basis. Unlike purposes, users will certainly know how they are using the data. However, there are no widely recognised or standard definitions of activities as there are for user types and purposes. This may result in a tension between validity/reliability and transferability/comparability. Surveys that achieve validity and reliability by describing activities in ways that are clear to potential users are likely to use descriptions of activities that are specific to that set of users and thus sacrifice transferability and comparability. For example, the activity of "measure market penetration" in the context of the marketing department of a commercial company may be clearly understood within that department. Users within that department may answer a survey question asking whether they are using open data for that activity with some confidence, creating high reliability and validity within the department. However, the same description may be meaningless for an academic user and thus it is not transferable. Surveys which use more generic descriptions (e.g. "benchmarking") are potentially transferable but users may interpret the descriptions of activities in inconsistent and unexpected ways reducing reliability and validity.

As with other user attributes, case studies and business population studies can overcome these tensions. Those running the studies can take the time to clarify what activities mean in different contexts and describe them in ways that are comparable across different contexts. But this can only be done at the cost of efficiency.

6.3 Impact

There are well-established attributes for economic impact at the macroeconomic level (e.g. growth, inflation, employment) and corresponding metrics (e.g. GDP, CPI, Percentage of workforce unemployed). For example, the paper Creating Value through Open Data uses four metrics:

- Market size and value added as a percentage of GDP;
- Number of jobs created;
- Cost savings for the public sector;
- Efficiency gains or productivity gains;

Macroeconomic studies are the only method that addresses the macroeconomic impact of open data. While these studies are very significant for assessing the impact of open data, they have limited application to specific portals, with the possible exception of very broad national portals such as the Spanish datos.gob.es, whose activities might reflect the levels and types of open data activities in the country as a whole. Given this limitation, macroeconomic studies, if done well, are valid and reliable but use a lot of resource and are therefore not efficient.

It is more plausible to measure the economic impact of a specific portal at the microeconomic level. The common attributes of interest at the microeconomic level are very similar to the possible purposes at the user level, e.g. profitability and sales. The key difference is that when considering impact, we wish to measure whether those purposes have been fulfilled. The metrics are therefore different in nature. For example, instead of a count of users seeking to increase profitability we want to measure how much profitability actually did increase.

While this could be done via a user survey, it is likely to suffer from very low validity and reliability. In most cases, it is extremely hard to assess the contribution of an open data portal to microeconomic objectives without detailed knowledge of how that data participates in the business processes. Acceptable levels of validity and reliability require a level of knowledge of the organisation that only come from a case study or a microeconomic study with corresponding inefficiency. If it is accepted that, despite the cost, a study of some kind is needed to measure microeconomic impact, there are significant benefits to measuring commonly accepted microeconomic attributes using a commonly accepted vocabulary of metrics. This contributes to transferability and comparability, and can contribute to efficiency through reuse of material in other studies. One such example of a study (and associated metrics) is the Spanish ASIDIE (the majority of these attributes are measured in Euros or counts of instances).

Attributes Measured by ASEDIE	
Subsectors of infomediary companies	Technical consulting, culture, directory service, economic and financial, publishing, market research, meteorological, geographic information, infomediation technology, tourism
Turnover	average, total, by subsector
Employee	total, by subsector, average turnover per employee, average expenditure per employee, average wage per employee
Share capital analysis	total, by subsector, average social capital
Profit and Loss	total, by subsector
Analysis of commercial risk	total, by subsector
Long term companies	sales evolution, employee evolution
Delisting	by motive (e.g. closure), community, subsector

Table 3 Attributes Measured by ASIDIE

6.4 Business Model

A business model can be treated as an attribute and the corresponding metric is simply to count how many users (where a user might be an organisation or an individual) of the portal fall into each business model. However, this presents formidable problems. There are many different proposed frameworks for open data business models (Lindman et al list four proposals in the literature prior to their own). Consequently, there is no widely accepted and understood vocabulary of business models. Furthermore, assigning an open data user to a business model requires considerable resources including expert, but subjective, assessments. Therefore, using a business model as a metric is likely to have low reliability and low efficiency. If the business model framework is clear, and described in sufficiently general terms, then it should be transferable across many situations (indeed this is one of the reasons for having a framework) but the low validity will limit comparability.

The table below summarises the information above, and assesses the methods using the criteria in Table 1. The three methods: case studies, population studies and microeconomic studies frequently have the same considerations and so they are collectively called “Studies”. Sensitivity is omitted from the criteria for assessing the methods because all the methods come to the same conclusion. If a method is valid (a prerequisite for being sensitive) then there no theoretical reasons to doubt the method’s sensitivity although this might change in practice.

Attribute	Plausible Method	Valid	Reliable	Efficient	Transferable	Comparable	Comment
DATA							
Dimensions	Not covered in this paper						
Sectors covered	Current automated methods	Good	Good	Very good	Good	Good	Requires metadata using standardised definitions. Established definitions are widespread.
Core data	Current automated methods	Good	Good	Very good	Good	Good	Requires metadata using standardised definitions. Several definitions exist.
Data quality	Current automated methods	Medium	Good	Very good	Good	Poor	Current automated methods are not relevant to some data quality attributes e.g. accuracy
	Studies	Good	Good	Poor	Medium	Medium	May be required for some data quality attributes such as accuracy.
USE							
Type of user (job title, industry, demographic)	User Survey	Medium	Medium	Good	Good	Depends on category	Can take advantage of widely recognised definitions.
	Studies	Good	Good	Poor	Medium	Medium	Poor efficiency prevents studies being used in an on-going fashion and limits them to snapshots of specific user groups.
Purpose (common business objectives e.g. increase revenue)	Survey	Low to medium	Low to medium	Good	Medium	Medium	Definitions are not as well established as type of user – but in many cases will still be recognised quite easily.
	Studies	Good	Good	Poor	Medium	Medium	Poor efficiency prevents studies being used in an on-going fashion and limits them to snapshots of specific user groups.
ActivitieS	Survey	Medium to good.	Medium to good.	Good	Medium to good.	Low to medium.	Lack of standard definitions or commonly understood terms leads to a tension between validity/reliability and transferability/ comparability.
	Studies	Good	Good	Poor	Medium	Medium	Poor efficiency prevents studies being used in an on-going fashion and limits them to snapshots of specific user groups.
IMPACT							
Macroeconomic	Study	Good	Good	Very Poor	Good	Good	If done well, macroeconomic studies can be valid and reliable but the cost is extremely high and are limited to broad national portals.

Attribute	Plausible Method	Valid	Reliable	Efficient	Transferable	Comparable	Comment
Microeconomic (common business objectives similar to user purpose)	Survey	Very low	Very low	Good	Medium	Medium	While a user survey could be designed that was efficient and comparable the low validity and reliability mean this approach has limited potential for measuring microeconomic success
	Study	Good	Good	Low	Good	Good	Poor efficiency limits the potential for repeating this method. However, use of commonly accepted attributes would greatly improve transferability and comparability.
BUSINESS MODEL							
Wide choice of models	Study	Low	Low	Low	Low to Good	Low to Good	Transferability and comparability may be good if the same models are used. However, this is of limited value given the low validity and reliability.

Table 4 Assessment of Methods and Metrics

7. Future Automated Methods

The potential for future automated methods is hard to assess. As discussed above, current automated methods can account for data/access attributes and future automated methods offer little additional value in this context. The potential benefit is in being able to measure attributes in the user and impact areas. While the EESODP report suggests future automated methods may throw light on user attributes, they are unlikely to be able to do so directly. For instance, data citation technology can potentially provide data on the users of open data sets – but if the method is to be automated, then the data about the user will be limited to what can be obtained automatically. This might include such things as the organisation the user works for, or the software using the data. It is unlikely to include the user's job description. This would have to be deduced from the automatically available data, and it is hard to see how this can be done validly and reliably without human judgement. The connection between the data that can be obtained automatically and the relevant attributes becomes even more tenuous when considering purpose and activity. An alternative is to include technology that asks users about role, purpose or activity when they make a data citation. This approach could be valid but is effectively a variant of the user survey method discussed above.

However, a possible automated method for use tracking is based in encouraging greater on-portal activity via community data spaces. Currently, with one or two notable exceptions, users are not specifically encouraged to engage with data portals in a meaningful way. In order to more effectively track use, it is key to develop portals in the direction of more collaborative environments where the

user is encouraged to engage with the portal (via other users) rather than extract the data and leave. Such an environment can be found in other data communities, such as those that use version control (VC), for instance, GitHub. Increased onsite activity would also mean the effort of finding links and improving data quality would be shared with data consumers, distributing the effort required to maintain and improve data quality among those benefiting from the data. As a side effect of using such technologies, data publishers would have access to more granular data on how their data is used, which in turn would allow them to identify high value datasets and ascertain the impact of open data. A further benefit of using VC is this community has already begun to consider the challenges around IoT data, for instance, how to manage extremely large files, such as 240 million rows of parking sensor data, and managing data aggregation. An example is found in Report 4.4.2 Leveraging Distributed Version Control Systems to Create Alternative Portals. It is likely such activities will become more relevant in cases where dataset search across multiple locations is implemented (e.g. Google Dataset Search).

As open data is published online this has allowed the utilisation of site and related analytics for the measurement of some open data activity. For instance:

- **Page analytics.** These are similar to other types of web site, which record metrics such as which pages are accessed most often and the order in which they accessed.
- **Downloads**. Which datasets are downloaded and how often?
- **API metrics.** Where the portal enables users to access data through an API it is possible to record how often the API is used and some data about who is using it.

These automated assessment metrics are the longest standing portal level indicators, but they are considered limited in their application for assessing use. However, research has demonstrated that, when combined algorithmically, they can be used to develop an accurate proxy for reuse. The report Automated Assessment of Indicators and Metrics describes a process for doing this and automatically assessing it

These kinds of future automated methods depend on significant cultural and organisation developments as well as technology. So, while future automated methods may be efficient and transferable in measuring user attributes, their validity, reliability, sensitivity and comparability are unknown at this stage and require further research.

The prospects for using future automated methods to measure attributes in the impact area seem quite small. Methods such as data citation take measurements at the time and location when the data is being used. This is desirable when measuring use. However, the impact of use may happen to a different group of people in a different location at a different time – in some cases months or even years later. Automatically linking the impact to the data appears to be a formidable task of dubious value.

Social media analysis has great potential for measuring both use and impact. It can call upon a large body of research and experience in analysing social media in multiple different ways for purposes ranging from measuring consumers' reaction to products to predicting voting intentions. Methods include: natural language processing, news analytics, opinion mining, scraping, sentiment analysis,

and text analytics [15]. This offers the possibility of measuring use and impact attributes for users other than primary users in an efficient and therefore ongoing manner. For example, it is quite conceivable that analysis of Twitter traffic that mentions a portal will give a valid and reliable measurement of the impact of that portal.

However, social media analysis has also had notable failures and there is very little research or experience in using social media analysis for measuring open data use. It remains to be discovered if social media analysis is useful in the context of open data, which analysis methods are the most effective for which attributes, and what the barriers and aids to using these methods in this context.

8. Recommendations

Once the attributes of interest are decided upon, the types of metrics required to assess them should be sourced from well-established statistical and data sources. This increases transferability, reliability and validity.

Ecosystem organisers, whether from the business or local government communities, can use these insights to engage with the right partners to enable reach to users who may not be directly engaging with portals. Business groups should be encouraged to survey their members for open data use and impact.

Publishers should aim to engage with reusers identified in show cases/use cases to develop quantifiable indicators.

Publishers and portal managers should share lists of metrics they have identified with other publishers and portal managers, in order to encourage larger and more comparable catalogues of metrics.

Attributes should be regularly checked to ensure that portal owners or publishers are measuring the right concept – use or impact – and not conflating the two.

9. Conclusion

Measuring the economic impact of open data portals remains a formidable task. It falls into two main subtasks: deciding what attributes to measure and identifying good methods for measuring those attributes.

CAF is a useful framework for the first task, but it is currently at a high level and does not take advantage of experience in measuring the impact of other types of information. Recognised attributes from other information sources such as demographic data provide an additional level of detail, which has been used successfully for many decades. Using established attributes from other information

[15] Batrinca, B. & Treleaven, P.C. AI & Soc (2015) 30: 89. https://doi.org/10.1007/s00146-014-0549-4

sources potentially increases the usefulness of open data by facilitating integration and comparison with other sources of data and information. The approach which we currently see as most valuable is to gather a range of potential microeconomic/business objective attributes, and understand which of these can be most universally applicable, and easily measured. The challenge here is then to identify the most appropriate methods for measurement, given the requirement for secondary input from (re)users.

The choice of current methods is dominated by the trade-off between efficiency and validity. For the most part, when measuring use and impact, methods that are efficient have limited validity. The decision on whether to emphasise efficiency or validity will depend on the objectives of the measurement exercise. For example, if the objective is to introduce a system of continuous improvement in an open data portal, then efficiency has to be a high priority, as measurement has to be ongoing and this is not feasible with a low efficiency method. Experience may teach the managers of the portal how to interpret the results better. On the other hand, if the objective is to justify continued funding for a portal, then a relatively resource intensive study that clearly measures use and impact may be a better choice.

Amongst current methods, the user survey is an attractive compromise for measuring user type and activity amongst primary users. In the future, new automated methods and social media analysis may avoid the need for compromise is some contexts. Future automated methods are unlikely to address impact, but offer a prospect of measuring user type and activity amongst all types of users. In theory, social media analysis has the potential to measure almost any aspect of use or impact amongst almost all types of users. However, very little is known about the practicalities of using social media analysis.

Whatever methods are being used, there are clear advantages to adopting consistent standards across the open data world and making those standards consistent with other relevant measurement programmes where practical. This increases efficiency by facilitating reuse and greatly enhances transferability and comparability of all methods. There is also significant potential in examining how methods can be usefully combined. For example:

- A microeconomic survey can be used to calibrate an automated method and thus increase the ongoing validity of the automated method.
- A user survey may be an efficient way of determining which user types should be the subject of a microeconomic study