Benchmarking open data efforts through indices and rankings: Assessing development and contexts of use

Martin Lnenicka, Mariusz Luterek, and Anastasija Nikiforova

Abstract: *This paper aims to provide a broad perspective on the development of benchmarking open data efforts through indices and rankings over the years, both at the level of countries and allowing for a cross-country comparison. The methodology follows a systematic search for the relevant resources, their classification and identification of six open data benchmarks to be further analyzed, the identification of their key components through decomposition, their description, and identifying the similarities and differences. Three major groups of indices and four periods that characterize the efforts to benchmark and measure the development of open data are identified, where the first measure the openness of the selected categories of data, the second focuses on different aspects of the open data ecosystem, using a large number of variables, and the third is a combination of both approaches. Recommendations as well as trends that can form the benchmarking frameworks in the future are also discussed. The findings are of a high importance for individual countries, which allow for correct and accurate interpretation of the results changes in the scope of a given index or rank, i.e., whether the difference in results is the result of national efforts or the subject of changes in the specific index, as well as how to combine and interpret the results of a number of indices for correct decision-making and for the definition of the future actions where the results vary significantly. In addition, the findings are also important for international organizations publishing benchmarking reports.*

Keywords: Open data; Benchmarking; Index; Ranking; Development; Indicator

1 Introduction

Benchmarking has its origins in the private sector, but public sector involvement was necessitated by the globalization process that tend to press for comparable outputs and standardized sets of policies, procedures, and practices. According to Rorissa et al. (2011), "*benchmarking is used as a tool for making comparisons between two or more entities based on a defined set of indicators."* The main force that drives globalization is the spread and penetration of Information and Communication Technologies (ICT) among businesses, citizens, and public sector agencies and institutions. E-readiness, ICT capacity, ICT preparedness, and ICT penetration measures were introduced to capture this development and allow comparisons between countries. After 2000, e-government development was the focus of these efforts. Various indices and rankings that assess the strengths and weaknesses or the overall progress of selected countries were published focusing on different points of view in delivering services and projects on the country or local levels. Over the years, the attention has shifted from requirements given by the type of demand and supply, front-end and back-end services to the needs of businesses and citizens and their engagement in the decision-making processes.

Open government has come to the forefront of wider public attention due to increasing pressure on public sector institutions to be more transparent and accountable to the public. The change in provision of public sector information was also enabled by the proliferation of ICT among the public. Open government builds on the democratic practices and modern ICT, i.e., e-democracy as a way of providing communication and cooperation channels closer to citizens and facilitate their wider engagement in public policy-making. E-government then serves as a connecting element which absorbs these emerging dimensions and, on that basis, can provide better services that meet the requirements of

citizens. This resulted in the development of open government initiatives that aim to disclose Open Government Data (OGD) through open data portals and other channels.

Although public data have been available online since the Internet has become widespread, especially official national and global statistics, the rise of OGD after 2010 was a key turning point in efforts to benchmark countries' transparency and openness-based policies and actions. If a research area is new or already established approaches do not apply, it takes a while to establish a robust methodology that provides relevant outputs and feedback. It is also needed to reflect changes that are associated with every relevant process that occur in the ecosystem and significantly contributes to the output (Bannister et al., 2007; Lämmerhirt and Brandusescu, 2019; Susha et al., 2005; Zuiderwijk et al., 2021). It is evident that through the years some of the processes will become less significant and must be removed from the benchmarking framework. In this regard, this paper aims to provide a broad perspective on the development of benchmarking open data efforts through indices and rankings over the years both at the level of countries and allowing for a cross-country comparison. It distinguishes between global reports published by international organizations and benchmarks found in the literature.

Considering the existence and broad use of a synonyms for terms related to the topic we are studying, let us define the key terms that we use consistently in this study. The term benchmark allows for a comparison of two or more countries using a framework to receive a level of score. This score is measured by a formula provided by a methodology, including measuring primary data in standard units. The term assess is used to decide the importance and development of open data efforts.

Governments sometimes assess or declare the development of OGD and OGD portals over the years, based on the same ranking, but rarely consider that changes may relate not only to the success or failure of efforts, but also to changes within a specific ranking system. Therefore, we are covering six ranking systems, more precisely the Global Open Data Index (GODI), Open Data Barometer (ODB), the Open, Useful and Re-usable data (OURdata) Index, the Open Data Inventory (ODIN), the Open Data Maturity Report (ODMR), and the Open Government Development Index (OGDI), during the years, which is identified as a result of a review of open (government) data related websites of international organizations, relevant reports, and search engines and a review of literature, focusing on the changes that take place in this system. This should provide value by allowing it to be determined whether the changes in the OGD and the corresponding portal relate only to their specific changes.

The paper is organized as follows: the following section describes the research methodology as well as the research questions. This is followed by sections dealing with research background. The next section includes the comparative analysis of global open data benchmarks, open data benchmarks in the literature, and the summary of the results found. It is followed by the discussion and limitations section. Finally, conclusions summarizing the lessons learned are provided.

2 Methodology

Under the above aim, the following research questions (RQ) were defined:

RQ1: What indices, rankings, and other frameworks are used to benchmark and measure open data efforts and what is their structure?

RQ2: What open data benchmarks are used in the literature?

RQ2.1: How often do researchers refer to open data-related indices and rankings in their studies?

RQ2.2: How often do researchers use open data-related indices as the focus of the study? And what is the main goal of these studies?

RQ2.3: What indices, rankings and other frameworks are the most popular among researchers, i.e. used to benchmark open data efforts, support their assumptions and / or findings?

RQ3: How these benchmarking tools evolved through the years and how they reflect new trends and modern ICT?

The RQ1 was addressed through analysis of selected open data benchmarking efforts, identified by search in the Google search engine, using the phrase "*open data index / ranking*". It led not only to finding indices and rankings with open data term in their titles, but also those which were referred to as an open data index / ranking by media, city officials and fellow researchers. The following criteria were used in the final selection of the indices and rankings for further analysis. Each ranking had to 1) have at least one index / ranking published, 2) include a methodological statement, 3) be published by international organizations at the level of countries and 4) have a global reach. Both reports fully dedicated to open data benchmarking or having a relevant section on open data were taken into consideration.

For the purpose of answering the RQ2 and its sub-questions, a systematic review of literature was conducted on five digital libraries – *ACM Digital Library*, *ScienceDirect*, *IEEE Explore*, *Springer* and *Emerald Insight*. The search was conducted querying on these data sources keywords "*open data*", "*open government data*" and "*OGD*", "*index*", "*indexes*", "*indices*", "*rank*", "*ranking*", "*rankings*". They were combined using Boolean operators AND and OR. First, we aim to find out how often do researchers refer to indices and rankings in their open data-related studies. Then, we look at how often indices or rankings are the primary focus of open data-related studies, and then we finally refer to the popularity of certain indices and rankings in these studies.

To answer the RQ3, we analyzed the classified and decomposed indices and rankings, including corresponding reports, identified trends, and compared them to find the similarities and differences, i.e., it was addressed through synthesis of results achieved while answering the RQ1 and RQ2.

3 Research background

As mentioned above, the effort to benchmark the development of open data is based on the concept of e-government. This also means that most benchmarking frameworks have their basis here. The question is whether this approach is the right one, because e-government is essentially based on the availability and quality of infrastructure and services. Open data, respectively open government and governance, however, deal with the processes of interaction and cooperation between stakeholders. While e-government is still a rather static conception, the open data ecosystem, in which these various processes take place, is a dynamic system requiring a different approach. Based on the literature review, it is possible to identify developments where some indices and rankings have been and still are part of e-government reports. Some later separated and, of course, most were created completely new.

Benchmarking of open data efforts must be addressed in the context of other elements of the system, and our view on it is shown in Figure 1. It is the link to e-government implemented on a country-level that significantly affects the open data ecosystem. Many countries still do not have policies and strategies that focus directly on open government and open data and are usually part of e-government strategies or those related to (ICT) digitization. For this reason, it is difficult to fully capture the impacts of open government and open data on society. The issue of stakeholders, their requirements,

and especially the knowledge and skills to participate in the open data ecosystem must be one of the key pillars to consider. When it comes to decision- and policy-making other concepts of e-services enter here, which are the output of the use of open data, such as informed decision-making in e-voting and e-elections.

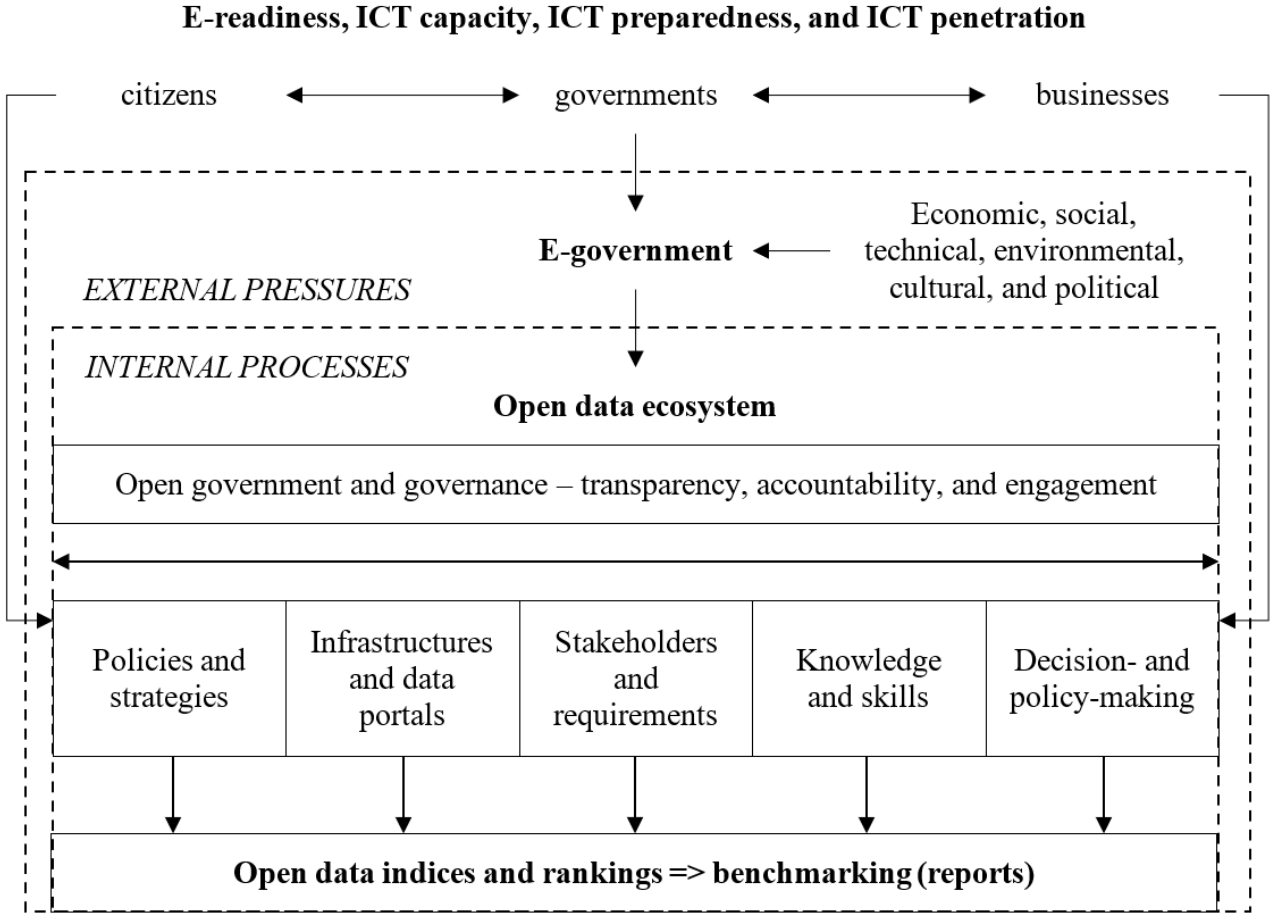

Figure 1: Benchmarking open data efforts – a system view.

3.1 E-government and ICT-related benchmarks

Before the focus on open government and open data efforts, but, at the same time, when the Internet and other digital technologies were already one of the main communication channels between the stakeholders, various indices and rankings included metrics related to ICT capacity, preparedness, and penetration. E-readiness models consisted of various dimensions, which included infrastructure and bandwidth, availability, quality and prices of connections, knowledge and skills of stakeholders to use ICT, maturity of online services etc. The concept of the information society has become more expansive since the 1990s and is continuously affecting information flows and the use of information regarding ICT trends. The key stakeholders who must consider the current development of this society, and at the same time are helping to shape it with their decisions and actions, are governments.

E-government is about building relationships between stakeholders that are efficient, effective, responsive, and transparent. For the last more than 20 years, e-government and ICT-related benchmarks have strongly established the approaches and structure of benchmarks in the public sector, including both positive and negative connotations (Luterek, 2020; Skargren, 2020). In this regard, this concept affects how open data efforts are benchmarked and what outputs and recommendations are expected from them. Among the most widely recognized e-government and ICT-related benchmarks are the United Nations Department of Economic and Social Affairs (UN-DESA) with the E-Government

Development Index (firstly published in 2001, the latest report is from 2020), the World Economic Forum (WEF) with the Network Readiness Index (first launched in 2002, the 2019 and 2020 editions are grounded on the Portulans Institute), and the International Telecommunication Union (ITU) with the ICT Development Index, which was published annually between 2009 and 2017 (Bannister, 2007; Máchová and Lněnička, 2015).

Most of these indices and rankings are composite in nature and comprise of various dimensions and / or metrics. Through the years, some of them were removed, their weights were increased or decreased, and some new dimensions and metrics were introduced. Thus, these changes that correspond to the evolving information society affect the frameworks of those indices and rankings as well as the methodology behind them. However, these changes sometimes encounter conflicting views regarding the economic, social, technical, environmental, cultural, and political pressures of involved stakeholders that disagree on what metrics describe the phenomenon the best and what methodology should be used (Luterek, 2020; Skargren, 2020). This is the case of the ICT Development Index which was lastly published in 2017 and since then there is no agreement on the improved methodology that will reflect ICT trends.

Disclosing Public Sector Information in the form of Open Government Data (OGD) is considered as one of the e-services of e-government. E-government policies and strategies strongly influence what is measured, i.e., the progress of policies implementation and identifying best practices. Some examples include Obama's open government strategy, EU's eGovernment Action Plan, which outlined that e-government requires a broad domain of ICT practices, including Key Enablers like open data. In this regard, e-government and ICT-related benchmarks that are usually accompanied with corresponding reports include chapters or overviews focusing on open data. If there is a demand for these outputs, area-specific indices and rankings are introduced. The UN-DESA provides the E-Participation Index, the Local Online Services Index, and the Open Government Development Index (OGDI). Therefore, open data indices and rankings are closely related to those dealing with e-government and usually share the same metrics in their frameworks.

3.2 Open data ecosystem

Building a robust, reliable, and comprehensive framework to benchmark and measure open data efforts is a real challenge (Welle Donker and van Loenen, 2017). An ecosystem approach is useful when trying to better understand the elements and their relationships that occur in the study area. It can be considered as a lever to mitigate the risks (Martin et al., 2017). According to Lněnička and Komárková (2019), this approach ensures that "*the public agencies and institutions will fulfil their tasks with quality, efficiency, and stakeholders' satisfaction, as required by initiatives of open government movement*." By identifying, addressing, and monitoring all the relationships, we can define all the elements in the ecosystem as well as their changes and development of requirements in time that should be addressed while updating the benchmarking framework.

The processes that drive the open data ecosystem result from the open government and governance concepts. All the relationships that exist between elements should support transparency, accountability, and engagement. The first element that usually defines these efforts is represented by corresponding policies and strategies. According to Welle Donker and van Loenen (2017), these should define "*the legal context, standards to facilitate data interoperability, and a stable and sustainable network for users of the data*." They are related to the e-government development and define the goals that should be achieved within the time frame and through strategic infrastructure and services projects. Data collected by the public agencies and institutions are the main source of OGD. To disclose them,

there are sequential data lifecycle phases and activities ensuring that these meet all the standards and requirements. Open data infrastructures represented by open data portals and other data repositories at country and local levels are designed to provide features that enable to work with open datasets. Most of the benchmarking frameworks focus on the quality of these infrastructures.

While the previous paragraph deals with the supply, the demand for open data is realized by citizens, businesses, and other stakeholders such as researchers, journalists etc. In the open data ecosystem, they can have different roles in which they interact with open datasets and stimulate the ecosystem. These are data user, data producer, data prosumer etc. (Lněnička and Komárková, 2019; Martin et al., 2017). Each stakeholder has also its own requirements on what the datasets or the data portals should meet. On the other hand, stakeholders are limited by their knowledge and skills. It is obvious that full potential and the value that can be gained from open datasets is heavily affected by this. Thus, it is important to consider these limitations while designing open data portals and other data repositories. If the main goal of the most benchmarks is to measure the impact that open data can have in the society, they must differentiate between the stakeholders, their knowledge and skills, as well as roles they participate in the ecosystem, and the relationship between supply and demand, i.e., both front-end and back-end processes.

Only when these elements and their relationships in the open data ecosystem will be addressed and implemented in the benchmarking framework, we can create a clear picture of what is the value of these data for involved stakeholders and how important are they for their engagement in decision- and policy-making processes (Lněnička and Komárková, 2019; Welle Donker and van Loenen, 2017). However, since there are specific country-related requirements and environments which affect the open data ecosystem, the framework should be widely discussed before its implementation. Finally, there are different kinds of data that will lead to different kinds of services. Thus, we can identify different kinds of ecosystem with different value potential of open data (Martin et al., 2017).

3.3 Open data indices and rankings

Although the assessment of achievements or fulfilment of goals can be defined in the context of a single entity at one time, most key findings and best practices are achieved when compared to others over time thereby determining the presence and degree of evolution of a specific aspect and the overall state of the artefact. In this regard, international organizations and non-governmental organizations introduce different frameworks and benchmarks that aim to provide an overview of open data efforts in selected countries and initiate discussion on best practices and recommendations. As presented above, these efforts are usually framed by the concept of e-government.

Benchmarks dealing with open data began to appear in early 2010s while at first, they were focused on meeting the principles of open data. Government or e-government portals and National Statistics Office (NSO) websites provided access to these data. As a next step, open data portals with corresponding features to work with datasets were launched and the attention of frameworks and benchmarks was redirected to reuse of these data and their impact. These are usually accompanied by a report. The reports are published annually or once every two years. Each of them describes and compares the level of achievements of open data within a specific country. Most of them have experts who assess open data initiatives or have the studied governments self-report, which is subsequently verified by experts (Lämmerhirt and Brandusescu, 2019; Susha et al., 2005; Zuiderwijk et al., 2021).

In 2013, first editions of two open data indices were published. The first one is the **GODI** by Open Knowledge Foundation (OKF) that follows the state of the government open datasets and how they meet a set of principles that define openness of data and content. The second one is the **ODB** by

World Wide Web Foundation (W3F), which aims to provide an overview of best practices around open data. In 2015, the pilot edition of the **OURdata Index** by Organisation for Economic Co-operation and Development (OECD), the **ODIN** by Open Data Watch (ODW), and the **ODMR** by European Union (EU) were firstly released. The OURdata Index measures government efforts to implement the G8 Open Data Charter based on the availability, accessibility, reuse of government data. The ODIN assesses the coverage and openness of key data categories and their adherence to open data standards. It assesses only data published on the official website of the National Statistics Offices (NSOs). The ODMR provides a support to European countries to improve their open data activities. The newest index introduced in 2020 by the UN-DESA is the **OGDI**. Since all these indices aim to assess the progress over time their methodology is continuously updated to meet current trends and requirements. In addition to the fact that goals and dimensions of these indices sometimes overlap, but in some cases vary widely, their coverage in the context of countries may vary significantly. Therefore, Figure 2 shows the number of countries included in each index release by year, while Table 1 gives a summary on these indices, i.e., their launch date, the date of the last report, as well as the total number of reports currently available and the number of countries in both the first and the last reports. As can be seen from the overview in Table 1, these benchmarks were initiated in or after year 2013.

Table 1: Overview of open data indices and rankings published by international organizations.

Title	**Publisher**	**First report**	**Last report**	**No. of reports**	**No. of countries covered**	
					First report	**Last report**
GODI	OKF	2013	2016	4	60	94
ODB	W3F	2013	2017	5	77	30
OURdata Index	OECD	2015	2019	3	30	33
ODIN	ODW	2015	2020	5	125	187
ODMR	EU	2015	2020	6	31	35
OGDI	UN	2020	2020	1	191	191

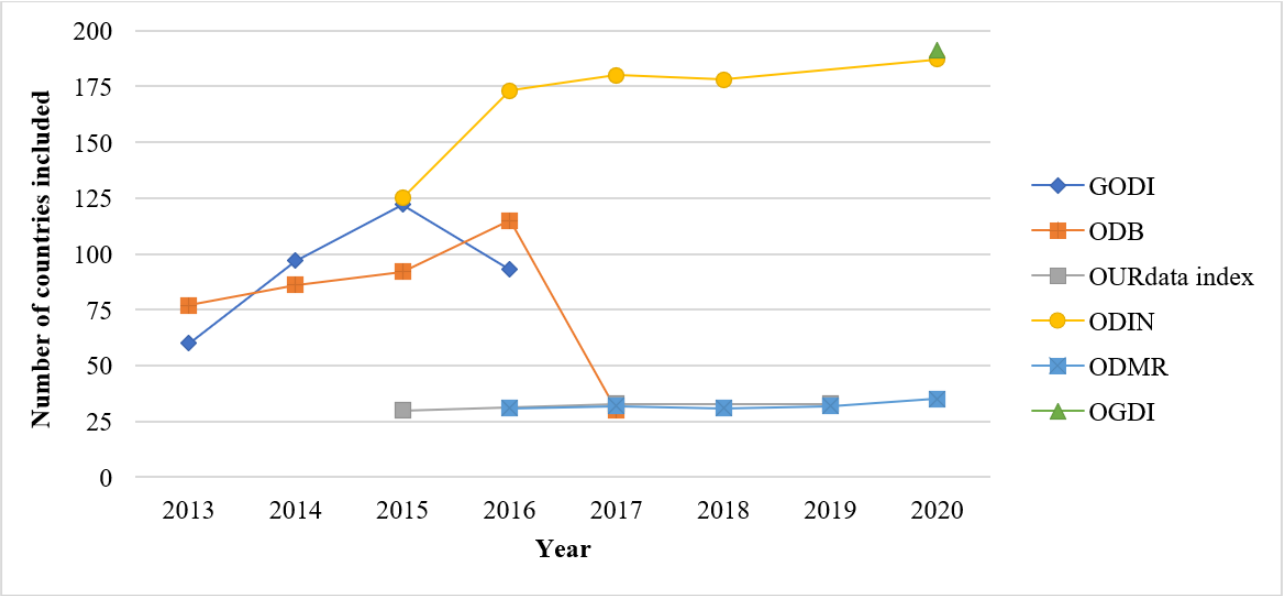

Figure 2: Number of countries included in the release of each index.

4 Comparative analysis and results

4.1 Global open data benchmarks

Although efforts are currently being made to bring open data closer to citizens at regional and local levels, typically Smart Cities, this trend is still new, so in this section we focus on benchmarks at the level of countries covering 6 benchmarks we have defined in previous section. This section answers the RQ1. The summary is also included in Section 4.3.

4.1.1 Global Open Data Index

The GODI was a series of studies conducted by the OKF between 2013 and 2017, with the last edition measuring the legal and technical openness of selected data categories through the fulfilment of six criteria, with the assumption that the dataset should be: openly licensed, in a machine-readable format, easily downloadable, up-to-date, publicly available, and free of charge. Those criteria were converted into questions and each of them received points that had different weights towards a total score through the years typically varying between 5 and 30 points per criteria and 35 to 65 points per dimension (Table 2). The index also ranks places and not countries, thus provides recommendations to sub-national governments that mainly operate autonomously from the higher national government.

The GODI aims to assess open data that has proven to be useful for the public. The first iteration of the report benchmarked only ten categories of the datasets: election results, company register, national map, spending, budget, legislation, statistical data, postcode/ZIP database, public transport database, and environmental data. The same method was used in 2014, with the coverage increased from 60 places from the previous report to 122 (OKF, 2014). The third report, published in 2015, used 13 data categories, from which four were newly introduced: procurement tenders, water quality, weather forecast, and land ownership. Additionally, the category public transport database was removed, and there was one label change – environmental data was renamed pollutant emissions (OKF, 2015).

The most recent report covers 94 countries and 15 data categories (OKF, 2017). As a result, the GODI benchmarks only some characteristics of the data publication and intentionally omits aspects like quality, impact, or usage. In this way, it is like the ODIN, although it covers more types of data providers. The categories are evaluated at the national level and cover the following topics: finances

and economy (budget, spending, procurement, company register), politics (election results), spatial data (land ownership, national maps, administrative boundaries, national postcode/ZIP database), law (draft legislation, national law), environment (air quality, water quality) and national statistics (OKF, 2017). The data collection process was based on the crowdsourcing approach and non-probability sampling technique. In this case, it was done through the work of the contributors who were identified as well-informed and knowledgeable in the context of governmental data in each location and could submit data for assessment (OKF, 2017).

Table 2: The method of scoring for the datasets in the GODI.

Dimension	Criterion	(OKF, 2014)	(OKF, 2015)	(OKF, 2017)
Technical openness	Does the data exist?	5	5	Not scored
	Is data in digital form?	5	5	Not scored
	Is the data available online?	5	5	15
	Is the data machine- readable?	15	15	20
	Available in bulk?	10	10	15
	Is the data provided on a timely and up to date basis?	10	10	15
	SUM	**50**	**50**	**65**
Legal openness	Openly licensed?	30	30	20
	Is the data available for free?	15	15	15
	Publicly available?	5	5	Not scored
	SUM	**50**	**50**	**35**

The overview of the GODI through the years can be found in Table 3. It shows two things: 1) when more countries are included, usually developing countries, the average score decreases, and 2) methodological changes, which usually follow the latest trends, again reduce average scores, as these new requirements are usually implemented more quickly by developed countries.

Table 3: A statistical overview of the GODI through the years.

Year	No. of countries	Change in methodology	Range of score [%]	Mean	Standard deviation	Minimum/ maximum
2013	60	NO	0–100	47.33	17.15	3/94
2014	97	NO	0–100	43.43	17.96	10/97
2015	122	YES	0–100	33.34	17.81	3/78
2016	93	YES	0–100	36.70	20.63	1/90

4.1.2 Open Data Barometer

The ODB report introduced by the W3F was firstly published in 2013 and the latest report from 2018 looks at how leading governments are performing a decade into the open data movement. The ODB was based on data collected through peer-reviewed expert survey responses, detailed assessments of the quality of fifteen types of datasets, and five secondary indicators published by other bodies. The 2016 report introduced an additional data collection tool in the form of a government self-assessment survey (W3F, 2016b).

The Barometer used three sub-indices, each built on three equally weighted components: readiness, weighted at 1/5 (government, entrepreneurs and business, citizens and civil society), implementation with the weight of the 3/5, (accountability, innovation, social policy), and impacts, weighted at 1/5

(political, economic, social), as shown in Table 4. The readiness sub-index was used to measure preparedness to secure positive outcomes from open data initiatives, which meant the existence of open data and mechanisms supporting engagement and reuse of data. The Government component benchmarked the existence of well-resourced open data initiatives on national and regional/city levels. It used pre-existing indices from the WEF's Global Information Technology Report 2013 (importance of ICT to government vision) and UN E-Government Survey 2012 (Online services index). In the case of "Entrepreneurs and businesses" component assessment included availability of data-literacy training programs for end-users and open data competitions supporting innovation and variables from the WEF's Global Competitiveness Index 2012 (firm-level technology absorption) and World Bank (Internet users per 100 people). Finally, Citizen and Civil Society component benchmarked the existence of well-functioning right-to-information law and regulatory framework for the protection of personal data, involvement of civil society representatives and information technology professionals in actions taken by the government regarding open data. Secondary sources used here were limited to Freedom House Political Freedoms and Civil Liberties Index (2013) (W3F, 2013).

The Implementation sub-index was built on experiences of the 2012 Web Index and assessed fourteen categories of data through detailed, ten-questions long checklists. The questionnaire was used to determine if the data exist and are available online, especially in machine-readable formats and in bulks, if it is free of charge, openly licensed, up to date, sustainable, and easy to find. For the Innovation component the list included data identified as of special interest to business: map data, public transport timetables, crime statistics and international trade data. Accountability cluster included data seen as especially important for government transparency: legislation, company register, national election results, detailed budget, and spending. The third component, Social Policy, benchmarked health sector performance, primary or secondary education performance data, national environment statistics, detailed census data, and land ownership data (W3F, 2013).

Online, mainstream media, and academic publications about open data were used to derive information relevant to the components of the third sub-index - impacts. The assumption behind this approach is that outcomes of the open data usage will bring journalists' and scientists' attention, which will result in measurable content. The questionnaire used in this section was focusing on the extent to which the open data have a noticeable impact on increasing government efficiency and effectiveness, increasing transparency and accountability in the country (which falls into the political impact category), environmental sustainability in the country, increasing social inclusion (social impact), and positive impact on the economy and building new businesses (economic impact) (W3F, 2013).

The 2014 Open Data Barometer introduced changes to the weights assigned to the three main sub-indices, increasing the importance of readiness and impacts (each assigned weight of 1/4) and limiting the role of implementation (2/4). There were also minor changes introduced to the components of the implementation sub-index: "innovation" includes "public contracts" from now on, and "land ownership data" was moved from "social policy" to "accountability" (W3F, 2015a).

The third edition of the report introduced another set of minor changes (W3F, 2016a). New weights were assigned to sub-indices: readiness – 35%, implementation – 35%, and impacts – 30%. The readiness consists of four equally weighted components from now on, which is the result of dividing the government pillar into two: government policies (existence of open data policy or strategy and consistent approach to data management and publication, supplemented by the Importance of ICT to government vision from the WEF's Global Information Technology Report) and government action (presence of well-resourced open government data initiative on central and regional or city levels, supplemented by UN's online services index) (W3F, 2015b). The same methodology was used in the

fourth edition of the Barometer (W3F, 2016b) and a special edition of the report from 2018, with a scope limited to 30 governments (W3F, 2018).

Table 4: Weights of sub-indices and components in the ODB.

Sub-index	Component	2013	2014	2015	2016	2017
Readiness	Government policies	1/3	1/3	1/4	1/4	1/4
	Government action			1/4	1/4	1/4
	Entrepreneurs & business	1/3	1/3	1/4	1/4	1/4
	Citizens & civil society	1/3	1/3	1/4	1/4	1/4
	Weight of the sub-index	**1/5**	**1/4**	**35%**	**35%**	**35%**
Implementation	Accountability	1/3	1/3	1/3	1/3	1/3
	Innovation	1/3	1/3	1/3	1/3	1/3
	Social policy	1/3	1/3	1/3	1/3	1/3
	Weight of the sub-index	**3/5**	**2/4**	**35%**	**35%**	**35%**
Impacts	Political	1/3	1/3	1/3	1/3	1/3
	Economic	1/3	1/3	1/3	1/3	1/3
	Social	1/3	1/3	1/3	1/3	1/3
	Weight of the sub-index	**1/5**	**1/4**	**30%**	**30%**	**30%**

The overview of the ODB through the years can be found in Table 5. The normalized score is rescaled to a 0–100 range between 2013 and 2016. Absolute values in the 0–100 scale for scores is used in 2018. The latest edition from 2018 includes only "leaders" and thus the mean score increases significantly compared to the previous edition of the rank.

Table 5: A statistical overview of the ODB through the years.

Year	No. of countries	Change in methodology	Range of score	Mean	Standard deviation	Minimum/ maximum
2013	77	NO	0–100	32.47	22.53	0/100
2014	86	NO	0–100	34.91	23.79	0/100
2015	92	YES	0–100	32.96	24.09	0/100
2016	115	NO	0–100	32.50	23.64	0/100
2018	30	NO	0–100	49.47	16.85	22/76

4.1.3 OURdata Index

The OURdata Index is used to benchmark the long-term sustainability of open data policies at the central level across OECD member states and partner countries (Pérez and Emilsson, 2020). The first edition of the OURdata Index was published in 2015 as a pilot study based on the G8 Open Data Charter (OECD, 2015), which defined a set of five principles for open government: Open Data by Default, Quality and Quantity, Useable by All, Releasing Data for Improved Governance and Releasing Data for Innovation (G8, 2013). The second iteration of the report has been adjusted to cover the principles of the International Open Data Charter (IODC), which makes the results from the pilot edition and those from the Index 2017 incomparable (Lafortune and Ubaldi, 2018). IODC introduced six principles for open data policies: Open by Default, Timely and Comprehensive, Accessible and Usable, Comparable and Interoperable, For Improved Governance and Citizen Engagement, For Inclusive Development and Innovation (ODC, 2015).

The first report introduced three dimensions of good practices in implementing open data solutions: data availability (provision of various data types on multiple subjects on national portal), data accessibility (publishing data in machine-readable formats on national portal), pro-active governmental support to enable innovative reuse of the data. All sub-dimensions were equally weighted at one-third (OECD, 2015).

The Data Availability dimension was benchmarked by verifying the presence of specific datasets on the national portal and applications reusing public data including topics such as national election results, national and local public expenditures, and the most recent national census. Similarly, the second dimension was used to verify the existence of specific features on the national portal as well as a general approach to data provision: notifications when new data are added, voting button, ranking popularity of data sets, systematic provision of metadata, and use of CSV format. Final dimension - government support to the reuse of data, included sub-indicators describing government dedication to fostering engagement of external stakeholders (regular consultations, software development contests, info session for citizens and businesses, co-creation events, data promotion to journalists) and creating data-provision friendly internal environment (data analytics teams in government, training for civil servants, OGD policies used as part of performance assessment). Data collection was done through a survey among central open data officials, as well as central portal analysis. All sub-indicators had a binary yes/no value assigned, with only CSV format having additional intermediate stage defined (OECD, 2015).

The new methodology, introduced in 2017, and used again in the 2019 report, was based on the same three pillars. However, instead of binary indices used in the previous iteration, it introduced sub-pillars built on data collected in an 80-questions survey and 140 data points. “Data availability” pillar describes policy on content open by default, stakeholder engagement for data release, and provision of various types of data on multiple subjects on the national portal (labelled “implementation”). The second dimension, “Data accessibility”, includes three sub-pillars: policy on unrestricted access to data, stakeholder engagement for data quality & completeness, and implementation (provision of data on the national portal, which are free of charge, with an open license, and in machine-readable formats). The final pillar, “Government support for data reuse”, covers such sub-dimensions as data promotion initiatives and partnerships, data literacy programs in government, and monitoring socio-economic impact (Lafortune and Ubaldi, 2018).

The overview of the OURdata Index through the years can be found in Table 6. Based on the statistical information we can say that, on the one hand, the average score is gradually improving although it should be stressed that release of 2015 cannot be really compared with 2017 and 2019, but above all, the countries that have had poor results are growing significantly.

Table 6: A statistical overview of the OURdata Index through the years.

Year	No. of countries	Change in methodology	Range of score [%]	Mean	Standard deviation	Minimum/ maximum
2015	30	NO	0–1	0.56	0.22	0.00/0.98
2017	33	YES	0–1	0.55	0.20	0.13/0.94
2019	33	NO	0–1	0.61	0.14	0.35/0.93

4.1.4 Open Data Inventory

Unlike the other benchmarks discussed in this paper, the ODIN is limited to official statistical data (Crowell and Swanson, 2020a). It includes 22 data categories and 65 statistical indicators to assess

two main aspects: coverage and openness. With the subject of the measurement limited to national statistical offices, their websites, and other connected websites and databases, the authors could assess standardized data, which allows the usage of a more detail-oriented methodology (Crowell and Swanson, 2020b).

The data for the assessment are collected by the team of assessors who search the relevant websites for data on the ODIN indicators and goes through several stages of the revision process to ensure high quality. The final score is based on the assessment of the data grouped in three main categories: social (10 subgroups), economic and financial (seven subgroups, and environmental statistics (five subgroups)), equally weighted. Each sub-group uses two to six representative indicators, which must be presented in specific categorical and geographical disaggregation. The ODIN Score is supplemented by two additional scores: Coverage Score and Openness Score (Crowell and Swanson, 2020b).

Coverage Score is calculated as the average of the scores in five dimensions, including scope, defined by the number of indicators and disaggregations, chronological range - data available for the last five and last ten years, geographical range - data available at the first and the second administrative level. Similarly, the Openness Score is the average score from five openness dimensions: machine readability; non-proprietary format; availability of bulk download, an API or user-defined download; completeness of metadata provided; and data with clear terms of use or open license (Crowell and Swanson, 2020b).

The general approach to the benchmarking process in the case of the ODIN remains unchanged since the first edition. The adjustments introduced in the consecutive iterations of the report are limited to the sub-categories, mostly at the representative indicators and country coverage level. The first ODIN report, published in 2015, assessed twenty sub-categories of national statistical data in 125 countries (ODW, 2015). The second report, published in 2016, covered 178 countries and introduced several improvements in the data collection process (Crowell, 2016). The 2017 iteration of the ODIN report added a new subgroup to the social category: Crime and Justice statistics, and covered 180 countries (Crowell, 2017), which number was decreased to 178 in the ODIN 2018/19 (Crowell, 2018). ODIN 2020/21 added one more subgroup to the social category: Food Security and Nutrition, and covered 187 countries in total (Crowell and Swanson, 2020b).

As can be seen in Table 7, the mean score of the ODIN is increasing through the years. Thus, it can be concluded that the openness of statistical data improves every year and at the same time contributes to the overall growth of the open data ecosystem, where citizens and other stakeholders can use different channels to search and work with open data.

Table 7: A statistical overview of the ODIN through the years.

Year	No. of countries	Change in methodology	Range of score [%]	Mean	Standard deviation	Minimum/ maximum
2015	125	NO	0–100	30.71	12.98	3.0/67.8
2016	173	YES	0–100	40.26	16.57	3.2/81.0
2017	180	YES	0–100	40.65	16.55	3.2/80.5
2018	178	YES	0–100	45.32	17.63	1.5/85.6
2020	187	YES	0–100	50.16	16.95	1.3/91.0

4.1.5 Open Data Maturity Report

In Europe, the European Data Portal (EDP) carries out an annual assessment of the national open data portals. The first report was published in 2015 but with no ranking. The ODMR, published annually

since 2015 by the European Data Portal, is presented as a landscaping exercise assessing the maturity of open data systems in the EU27, EFTA, and Eastern European Partnership countries (Knippenberg, 2020b). As a benchmarking tool, ODMR was created to play an important role in implementing open data policies across Europe, understanding progress, identifying bottlenecks and best practices.

The measuring method used in the ODR changes every year, which makes progress assessment difficult. The biggest update to the methodology carried out in 2018 creates two, hardly comparable, sets of data: 2015-2017 and 2018-2020. However, the data collecting process's approach remains the same and relies heavily on self-assessment surveys and desk research (Knippenberg, 2020b).

The first assessment of open data maturity was built on two pillars: readiness (up to 1000 points in five categories: the presence of open data policy; licensing norms; extent of coordination at national level; use of data; and impact of open data) and maturity (up to 250 points in three categories usability of the portal; re-usability of data; spread of data across domains) (Carrara et al., 2015). The scoring system favored the existence of proper policies on usage and re-usage of public data (maximum of 300 points) and actual usage (200 points), as well as the political, economic and social impact of open data (300 points). The second assessment, published in 2016, introduced only slight changes to the maximum points awarded in: open data policy (increased by 30 points), use of open data (increased by 60 points), the usability of the portal (decreased by 40 points) and re-usability of the portal (increased by 40 points), which resulted with the maximum final score increase from 1250 to 1340 points (Carrara et al., 2016). The method was then reused in 2017, with slight enhancement to the scoring, especially in favour of open data policy indicator (another increase by 70 points) and open data usage (increase by 40 points). Additionally, licensing norms and coordination at the national level were awarded 10 points more each, and the usability portal indicator was assigned 90 points, bringing it almost to its value from 2015 (Carrara et al., 2017). Changes applied in these three years prove that the ranking authors were looking for the best approach to benchmark open data implementation with a consistently growing emphasis on the policy framework and data usage.

The new methodology, introduced in 2018, is based on four dimensions, covering open data policy, portals, impact, and quality (Radu and Cecconi, 2018b). The first three dimensions can be referenced to sub-indicators used in the previous reports, while the quality dimension was only addressed in additional, non-scored questions in the 2017 survey (Carrara et al., 2017). This strong correlation between indicators used in both methodologies proves that the change in the benchmarking method was not driven by the need to address different aspects of the open data phenomenon. Instead, it seems that the goal was to increase the efficiency of the method and reorganize the composition of the indicator to make it more readable.

Policy dimension, with a weight of 27%, included three subgroups: policy framework (previously "open data policy") with the maximum score of 180 points (previously 400 points), licensing norms – maximum score of 150 points (almost doubled in comparison to 2017) and coordination at the national level with a maximum score of 350 points (previously 140 points). It is used to describe the policy framework, long-term strategic thinking behind open data activities, and the level of support that local/regional authorities have from the central government in implementing open data solutions (Radu and Cecconi, 2018a).

The impact dimension has been significantly expanded to include not only political (130 points), social (110 points), and economic (130 points) context, but also strategic awareness (200 points) and environmental issues (80 points) (Radu and Cecconi, 2018a). In the case of political, environmental, and social subgroups, the goal was to verify if multiple examples of use and reuse of public data in those fields can be provided and if it is a subject of systematic monitoring. A different approach was

used to benchmark economic context where verification was limited to the existence of multiple studies on the macro and microeconomic impact of open data conducted by the central government. Finally, strategic awareness, used to assess if proper monitoring and measurement mechanisms of open data reuse and impact are implemented, is in fact, a provision-side metric.

The third dimension refers to the national portal and consists of four indicators: 1) Portal features (250 points), which is a combination of usability and re-usability indicators from the previous methodology; 2) Data provision (160 points), which is a continuation of previous "spread of data across domains" indicator; 3) Portal sustainability (120 points), referring to sustainable funding mechanisms in place and regular user satisfaction surveys; and 4) Portal usage (120 points), describing the usage of analytic tools for better understanding visitor's profiles (Radu and Cecconi, 2018a).

Finally, the newly introduced quality dimension was based on three indicators: automation (of metadata management on the portal, 100 points), data and metadata currency (measuring regularity of data and metadata updates as well as the availability of current and historical data, 210 points) and DCAT-AP compliance (benchmarking quality of metadata through compliance with DCAT-AP standard and supporting publishers in increasing the quality of the metadata they provide, 210 points) (Radu and Cecconi, 2018a).

The methodology used in the 2019 report introduced several major changes to how policy and quality are measured, minor scoring changes to indicators from impact and portal dimensions, and equal weights for all dimensions (Blank, 2019b). In the policy dimension, the update included an increase of the maximum score for open data policy by 30 points, removal of the licensing norms indicator, and addition of the "open data implementation" (Blank, 2019a). This new indicator monitors how effective the open data system is in providing guidelines to assist publication, publication plans, limiting costs of access, training civil servants, and providing access to and information on local sources.

The most notable changes of the maximum score awarded within the impact dimension were environmental impact – increase from 80 to 150, and strategic awareness – decrease from 200 to 140 points. Portal dimension was also re-adjusted, especially in the case of the spread of the data across domains indicator (decrease by 50 points), portal sustainability, and portal usage (increased by 30 points each) (Blank, 2019a).

The Quality dimension was almost totally reworked: with the removal of automation indicator, the introduction of monitoring and measures (150 points) and deployment quality and linked data (150 points), change of scoring for the two remaining indicators: currency and completeness (decrease by 60 points, previously - data and metadata currency) and DCAT-AP compliance (decrease by 30 points), it limits the comparability of results obtained in 2018 and 2019 (Blank, 2019a).

The 2020 Open Data Maturity Report maintained the same methodology used the year before, with one minor change in scoring, awarding additional 5 points to the maximum value obtainable in "coordination at national level" criteria within the policy dimension (Knippenberg, 2020a). With changes regularly introduced since 2016, the stability shown in 2020 gives reliable possibilities of comparing data presented in consecutive reports and monitoring progress.

The structure and scoring of points through the years is in Table 8, while a statistical overview can be found in Table 9. Also, in the case of this benchmark, the maturity of OGD is increasing.

Table 8: Scoring of dimensions and indicators in the ODMR.

Dimension	Indicator	2015	2016	2017	2018	2019	2020
Policy	Open data policy	300	330	400	180	220	220
	Licensing norms	70	70	80	150		
	Coordination at national level	130	130	140	350	215	220
	Open data implementation					210	210
	Weight				**27%**	**25%**	**25%**
Impact	Political impact	120	120	120	130	130	130
	Strategic awareness				200	140	140
	Environmental impact				80	150	150
	Social impact	60	60	60	110	120	120
	Economic impact	120	120	120	130	110	110
	Weight				**26%**	**25%**	**25%**
Portal	Usability of the portal	100	60	90	250	240	240
	Re-usability of the portal	100	140	140			
	Spread of data across domains	50	50	50	160	110	110
	Portal sustainability				120	150	150
	Portal usage	200	260	300	120	150	150
	Weight				**26%**	**25%**	**25%**
Quality	Automation				100		
	Data and metadata currency				210	150	150
	DCAT-AP compliance				210	180	180
	Monitoring and measures					150	150
	Deployment quality and linked data					170	170
	Weight				**21%**	**25%**	**25%**
Sum		1250	1340	1500	2500	2595	2600

Table 9: A statistical overview of the ODMR through the years.

Year	No. of countries	Change in methodology	Range of score [%]	Mean	Standard deviation	Minimum/ maximum
2016	31	YES	0–1	0.57	0.21	0.00/0.91
2017	32	YES	0–1	0.68	0.20	0.00/0.96
2018	31	YES	0–1	0.60	0.20	0.01/0.88
2019	32	NO	0–1	0.62	0.20	0.06/0.91
2020	35	NO	0–1	0.70	0.23	0.10/0.96

4.1.6 Open Government Development Index

The OGDI was conceptualized as a supplementary index to the Online Services Index, which constituted one-third of the e-Government Development Index and was first published in 2020. The fact that the OGDI is not an independent benchmark explains why the methodology behind it is not properly described, and it is not possible to understand how exactly it is calculated. It is based on data referring to three dimensions of open data: 1) policy and institutional framework, 2) national portal, 3) data availability in six domains (health, education, employment, social security, environment, justice), and promotion of application through various events (Aquaro, 2020).

The UN e-Government Survey includes various aspects of open data provision since 2014 through mostly binary indices, describing countries: with open data portals, providing data in machine-readable formats, offering open government data related legislation, publishing datasets in open standards, or providing guidance for using open data through online tools (Zhu and Barthélemy, 2016). The first pilot computation of the OGDI was done using data from the 2018 UN e-Government Survey, which has assessed 110 areas of providing public information and services for the Online Services Index. Seven of those areas were referring to the provision of open data: 1) information about using open data sets, 2) existence of open data on education, employment, environment, health and social welfare and labor, 3) existence of open data competitions, 4) existence of an open government data policy online, 5) existence of a national portal, 6) existence of data dictionary or metadata repository in the portal, and 7) ability to request new open data sets (Aquaro et al., 2018). The number of the areas benchmarked in the 2020 report was reduced to six with the removal of the "ability to request new open datasets" (Aquaro, 2020). The framework is discussed in further detail in Zheng et al. (2020). A statistical overview for the OGDI is in Table 10.

Table 10: A statistical overview of the OGDI.

Year	No. of countries	Change in methodology	Range of score [%]	Mean	Standard deviation	Minimum/ maximum
2020	191	NO	0–1	0.59	0.35	0/1

4.2 Open data benchmarks in the literature

The section deals with the RQ2 and the importance of this topic in the literature. For the purpose of the study, the search was conducted querying on five digital libraries, more precisely *ACM Digital Library, ScienceDirect, IEEE Explore, Springer and Emerald Insight,* previously defined keywords "*open data*", "*open government data*" and "*OGD*", "*index*", "*indexes*", "*indices*", "*rank*", "*ranking*", "*rankings*". They were combined using Boolean operators AND and OR. Only papers written in English were addressed. In terms of the scope, both journal articles, conference papers, and chapters were studied. For the period covered by these searches, we covered the period 2001–2021 to gain insight into the popularity trends of these topics over the years and to select the most up-to-date studies to be further analyzed. For the starting date, it was selected by looking at the oldest returned results.

First, we intend to understand how often do researchers refer to open data-related indices and rankings in their studies? (RQ2.1). For this question, two queries have been used: **Query I** ["open data" OR "open government data" OR "OGD"] and **Query II** [["open data" OR "open government data" OR "OGD"] AND ["index" OR "indexes" OR "indices" OR "rank" OR "ranking" OR "rankings"]]. The results of both queries are provided in Table 11 and Figure 3. The figure shows that Springer Link provides more both open data-related studies and open data-related studies covering indexes. Overall, 47.33% of studies covering open data, mention at least one index with the highest number of articles in Emerald Insight, where 67.52% of studies on open data cover indices.

Table 11: The numbers of results found in selected digital libraries.

Digital Library	Query I	Query II
ACM Digital Library	4283	2096
IEEE Xplore	8711	3609
Science Direct	12946	5832

Emerald Insight	21118	10453
Springer	1413	954
TOTAL	**48471**	**22944**

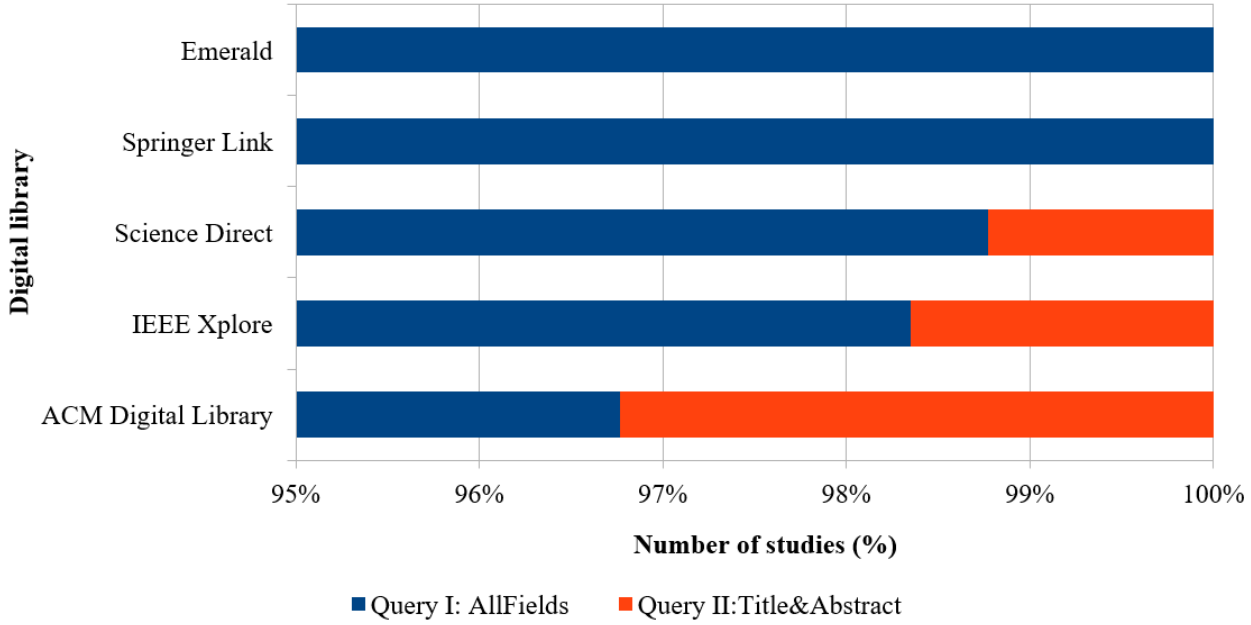

Figure 3: Distribution of studies by the element containing keywords.

For the purpose of the RQ2.2, i.e., *how often do researchers use open data-related indices as the focus of the study?*, the search was conducted covering only the second query - *[[“open data” OR “open government data” OR “OGD”] AND [“index” OR “indexes” OR “indices” OR “rank” OR “ranking” OR “rankings”]]*. To find relevant articles, dealing with rankings not only mentioning, querying of digital libraries was done in two ways: (1) abstract and title only, if the digital library allowed such queries, (2) anywhere in the text, i.e., including those that only mention them. The results of the analysis by digital library are provided together with the queries in Table 12. Figure 4, however, provides a more detailed insight, providing a summary not only by the digital library, but also by year.

One library, namely Springer, has not allowed to filter only those articles in which keywords are mentioned in both title or abstract, allowing them to be filtered out only by title, so its results cannot be considered comparable to other libraries. For those libraries that admitted to doing so, a total of 220 articles were found. In general, when all fields are processed looking for keywords used, 68611 articles have been found. It means the indices serves as the focus of open data-related studies in just 0,32% studies compared with the abovementioned 47.33%, where the indices are only mentioned. Therefore, it can be concluded that indices are more complementary materials that allow researchers to base and support their assumptions and / or findings rather build studies around them.

Table 12: The results of the analysis of digital libraries.

Digital Library	Query	Number of results (title & abstract)	Number of results (anywhere)
ACM Digital Library	*Title:(["open data" OR "open government data" OR "OGD"]) AND Title:(["index" OR "indexes" OR "indices" OR "rank" OR "ranking" OR "rankings"]) OR Abstract:(["open data" OR "open government data" OR "OGD"]) AND Abstract:(["index" OR "indexes" OR "indices" OR "rank" OR "ranking" OR "rankings"])*	70	2096
IEEE Xplore	*("Abstract":"open data" OR "Abstract":"OGD" OR "Abstract":"open government data") AND ("Abstract":"index" OR "Abstract":"indexes" OR "Abstract":"indices" OR "Abstract":"rank" OR "Abstract":"ranking" OR "Abstract":"rankings") OR ("Publication Title":"open data" OR "Publication Title":"OGD" OR "Publication Title":"open government data") AND ("Publication Title":"index" OR "Publication Title":"indexes" OR "Publication Title":"indices" OR "Publication Title":"rank" OR "Publication Title":"ranking" OR "Publication Title":"rankings")*	55	3609
Science Direct	*("open data" OR "open government data" OR "OGD") AND ("index" OR "indexes" OR "indices" OR "rank" OR "ranking" OR "rankings") in "Title, abstract or author-specified keywords "*	84	6739
Emerald Insight	*((title:"open government data") OR (title:"OGD") OR (title:"open data") OR (abstract:"open government data") OR (abstract:"OGD") OR (abstract:"open data")) AND ((title:"index") OR (title:"indices") OR (title:"rank") OR (title:"ranking") OR (title:"ranking") OR (abstract:"index") OR (abstract:"indices") OR (abstract:"rank") OR (abstract:"ranking") OR (abstract:"ranking"))*	11	954
Springer	*("open data" OR "open government data" OR "OGD") AND ("index" OR "indexes" OR "indices" OR "rank" OR "ranking" OR "rankings") in "Title"*	0	55213
TOTAL	-	**220**	**68611**

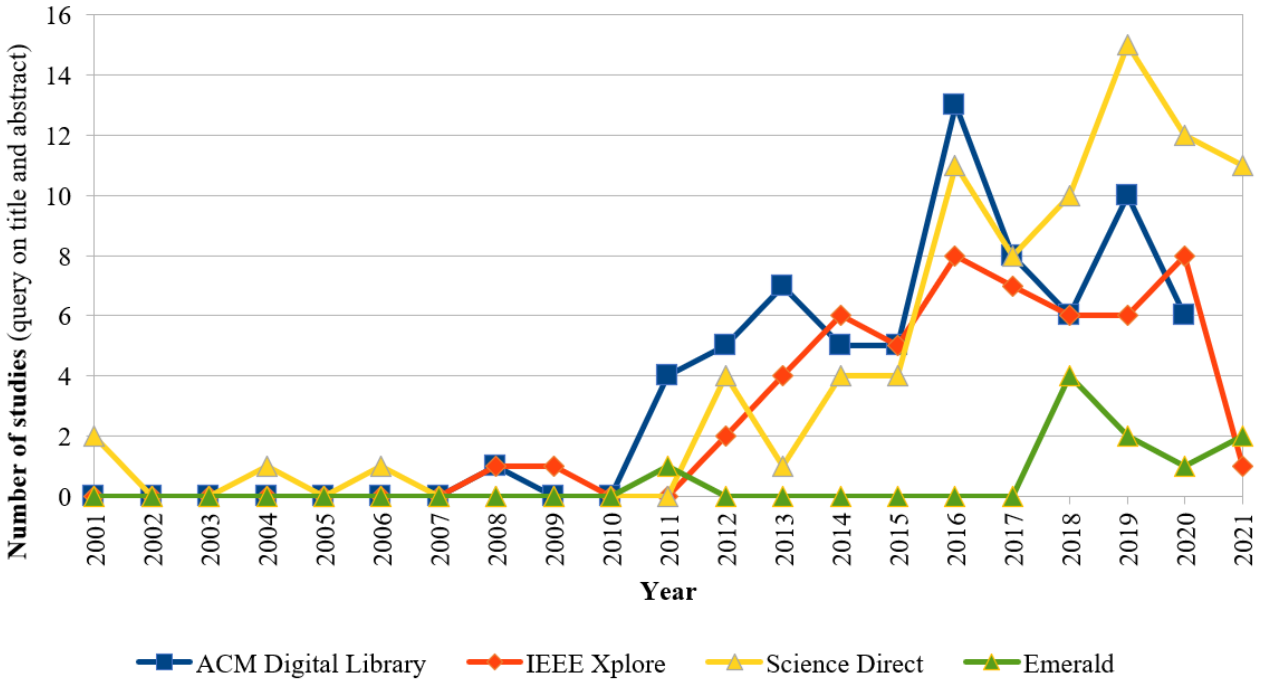

Figure 4: Distribution of studies by their number and year (2001–2021) (based on ACM Digital Library, Science Direct, IEEE Xplore, Emerald Insight).

Figure 4 presents numbers of studies that contain keywords of interest in a title or abstract conducted between 2001 and 2021 by a digital library, where only 4 of 5 digital libraries are presented. The excluded library in this Figure is *Springer*, i.e., digital library that has not provided an opportunity to filter studies that had search elements in abstract OR title. Between 2001 and 2010, indices were rarely mentioned or used as a central object for studies. This is related to the fact that there were only a few studies at these times which is consistent with our findings provided in previous Sections. However, from 2011, the indices and rankings became more popular and, as a result, they have started to be used in more studies.

In Figure 4 we observe several peaks of interest dated 2013–2014, 2016 and 2019. We assume that these peaks can be linked to the dates of the launch of European or / and global indices and rankings, since the GODI and ODB were launched in 2013, where the first peak was observed, but 2 years later three more indices have been launched, along with those previously launched and regularly maintained. Given that it is rather an assumption, we turn to an index-specific analysis to support or reject it.

Then, an analysis of the popularity of six indices in scientific literature, covering five digital libraries, have been conducted to answer RQ2.3, i.e., *what indices, rankings and other frameworks are the most popular among researchers, i.e. used to benchmark open data efforts, support their assumptions and / or findings?*. Given that the first indices, namely the GODI and ODB, have been released in 2013, this year has been chosen as a start year to be covered. The results of this analysis along with a trend-line set for every index are provided in Figure 5, where a summary of this analysis is provided for five libraries together. ODB may be seen as an undoubted leader among the analyzed indices, which are used more frequently compared with other indices, even now when it is not updated since the last report was published in 2017, with a slight decrease in its popularity starting in 2018. The next most popular index is the GODI, which was also launched in 2013 and has not also been updated for a while, more precisely since 2016. Despite this fact, it is the second most popular index to be mentioned in open data-related studies. It can be also seen that significant changes in terms of the

popularity of indices in scientific literature have been observed between 2017 and 2018, where the total number of studies covering the indices has increased by 44.94%, where this positive trend has been observed mainly due to the GODI and ODB, despite GODI has not been updated since 2016. In other words, indices are of high interest for researches even when they are not up-to-date and provide rather historical data. But for what purposes are these indices used?

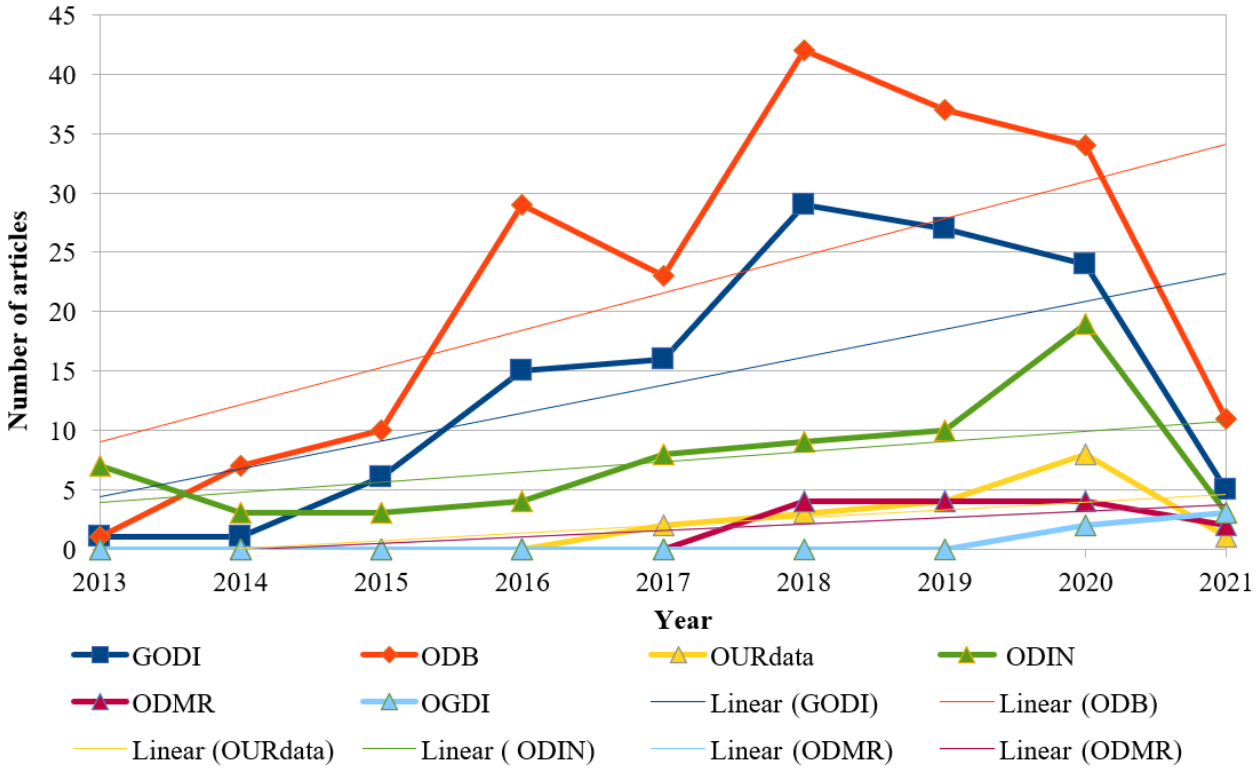

Figure 5: Popularity of indices in scientific literature between 2013 and 2021 (by IEEE Xplore, Science Direct, Springer Link, Emerald Insight, ACM Digital Library).

As a result of the initial literature analysis, we have selected the most relevant studies for their further examination. By relevant studies we mean studies that have a direct relationship with the open data index / ranking. This led us to 50 studies covering open data indices and / or rankings. These studies were inspected in three ways – (1) their main objective, (2) index and / or rank mentioned or used, (3) type of the study in relation to the extent of use of the "index" term. The last category requires categorization of these studies, depending on the way they are handled with the index. Two main categories have been involved – (1) background / state of the art / discussion, i.e., index or rank is mentioned in the relevant section and is rather supportive material than paper focus, (2) methodology, i.e., the index form a basis of paper and serves as input data of primary or secondary type. This classification is based on our subjective observation that, in most cases, indices and rankings are used in one of these two cases where the use of index or ranking to support or justify the results of a study or to introduce key concepts of a study and establish background should prevail. Each paper may, depending on its nature, be classified in one or both categories.

Table 13 presents the result of our analysis providing a type and summary of these studies and the list of indices / rankings extracted from them, according to which we can see that our assumption of a significantly higher popularity of studies using index data to justify their findings or establish background, rather than to establish methodology, is rejected. The number of studies in which index data serve as input is high and the nature of these studies may vary by using them as both primary or

secondary data. For the most popular indices to be used in these selected studies, the GODI, ODB, ODMR, and OURdata Index were found as the most popular and sometimes used in combination while developing another ranking system. The least popular are the Open Data Readiness Assessment (ODRA), ODIN, OGDI, Public Sector Information (PSI) Scoreboard and the World Justice Project (WJP) Open Government Index (OGI).

Table 13: Analysis of selected studies based on their type and index / rank used.

Source	Type	Brief description	Index / rank
Aarshi et al. (2018)	Methodology (input data)	The paper describes the various challenges facing Asian countries in adopting, promoting, and accepting OGD and OGD portals.	GODI
Ahonen (2018)	Background / state of the art / Discussion	The paper introduces a framework to examine changes in e-governance and applies it to a country case study on opening government location data. The framework consists of four dimensions: influential historical, social and political mechanisms; the diffusion of innovations; stages of e-governance development; and facilitators of change.	GODI, ODB, OURdata Index
Alarabiat et al. (2018)	Background / state of the art / Discussion & Methodology (input data)	An exploratory study aimed at investigating current EGOV assessment initiatives. The authors carry out it based on data from a desktop research and the global questionnaire addressed to 193 countries.	GODI, ODB, ODRA, OURdata Index,
AlRushaid and Saudagar (2016)	Methodology (input data)	The paper proposes a model to measure data openness level, which is based on a GODI and is used to measure data openness level of Saudi Arabian e-Government Data Portal.	GODI
Altayar (2018)	Background / state of the art / Discussion	The paper explores motivation factors for the adoption of OGD between public authorities in Saudi Arabia. A qualitative approach has been used to manage the study and the data have been collected using interviews and documentation	ODB
Altayar (2018)	Methodology (input data)	The paper studies the impact of government data openness on a knowledge-based economy at governmental level.	ODB
Amalia and Susanto (2019)	Background / state of the art / Discussion (partly Methodology (input data))	The paper proposes conceptual model for the analysis of motivation factors and perceived risk factors for the assessment of the open data in Indonesian local government.	GODI, ODB, ODIN, ODMR, OURdata Index,
Attard et al. (2015)	Background / state of the art / Discussion & Methodology (input data)	A systematic survey aimed at evaluating OGD initiatives. The OGD lifecycle and data consuming processes required within OGD initiatives, current approaches to such initiatives and their classification are proposed. Several assessments are discussed, and the list of challenges and issues are identified.	GODI, PSI Scoreboard, (+US Open Government Directive)
Bonina and Eaton (2020)	Background / state of the art / Discussion	Literature analysis to explore how service innovations are cultivated in the context of the OGD. A theoretical model has been proposed explaining how the owner of the OGD platform can manage the demand and the supply of the platform to promote the cultivation of the platform ecosystem.	GODI, ODB
Corrêa et al. (2017)	Background / state of the art / Discussion	The paper identifies how the current Brazilian OGD portals comply with the OGD principles.	GODI
Dawes et al. (2016)	Background / state of the art / Discussion	The paper proposes an initial ecosystem model for the planning and designing OGD programs by using the theory of social engineering systems and a review of existing open data-related research and practice guidelines.	GODI, ODB, OECD open data framework (OURdata Index since 2018), Open Data 500, Open Data census, Open Data Monitor, PSI Scoreboard
de Juana-Espinosa and Luján-Mora (2019)	Background / state of the art / Discussion (partly Methodology (input data))	The paper examines the development of OGD portals in the 28 EU countries. The paper provides a longitudinal and multidisciplinary perspective. The results show that EU countries are gradually homogenizing their OGD approaches in two “currents”, mainly based on economic factors and open government development status.	GODI, ODB, ODMR
Escobar et al. (2020)	Background / state of the art / Discussion	The paper proposes an approach of publishing statistical data from public repositories using Semantic Web standards such as RDF and SPARQL to facilitate analysis of multidimensional models.	ODB
Galeone and	Background / state	An overview of open data and their application in the social and health	GODI

Bonzi (2020)	of the art / Discussion & Methodology (input data)	care fields. The authors discuss the state of play in the world, paying particular attention to the situation Italian in relation to open health data.	
González-Gallego and Nieto-Torrejón (2021)	Methodology (input data)	The paper inspects whether the degree of openness and the coverage of datasets released by European governments have a significant impact on citizen's trust in public authorities.	ODIN
Grzenda and Legierski (2021)	Background / state of the art / Discussion	Propose a methodology for collecting and analyzing access data, describing the use of open data resources in individual applications. The methodology includes categorization of data collected from the portal, providing access to the underlying open data portals and third-party services.	GODI
Jetzek (2016)	Background / state of the art / Discussion	Danish case study intended to facilitate understanding of the problems of providing open access to government data through open data infrastructure. It aims to improve the quality of selected government data, make them more harmonized and improve their accessibility through the implementation of a common data dissemination platform.	ODB
Kawashita et al. (2020)	Methodology (input data)	Authors explore the different dimensions that OGD benchmarks assess to see what their measured aspects of publishing and using open data are. This is done by analyzing previous studies on how the principles of the Open Data Charter are assessed in the OGD assessments and supplying it with additional dimensions. The findings show that scope or focus of these benchmarks may differ significantly.	GODI, ODB, ODIN, ODMR, OURdata Index, Open Data Monitor
Kim (2019)	Background / state of the art / Discussion	Analyses the standard terms used in public data. The results of the study show that standard vocabularies created by the government need updates to reflect the nature of public data, and the relevant legislation and guidelines need to be revised.	GODI, ODB
Krotova et al. (2020)	Background / state of the art / Discussion	The authors compare the characteristics of open government and open company data to identify the necessary framework conditions for data sharing. They found that promoting legal certainty and economic impact is important policy measures to encourage data sharing.	OURdata Index
Kubler et al. (2018)	Background / state of the art / Discussion	The paper proposes an Open Data Portal Quality framework called ODPQ that allows users to easily rank open data portals in real-time.	eGovOI index, GODI, ODB
Máchová (2017)	Methodology (input data)	The paper examines the potential impact of open data on the level of corruption.	GODI, ODB, OURdata Index, PSI Scoreboard
Máchová and Lněnička (2015)	Background / state of the art / Discussion & Methodology (input data)	The paper incorporates open data into the e-government development benchmarking framework. It also analyses and compares the main prerequisites for e-government development, the benefits and risks of cloud computing, open data, including the concept of open big data, and participation tools in the public sector.	GODI, ODB, OURdata Index
Máchová and Lněnička (2017)	Methodology (input data)	The paper introduces a benchmarking framework to assess the quality of open data portals at national level. The results of its validation show that the number of datasets online and the complexity of open data portals and their functions differ.	GODI, ODB, OURdata Index
Marković and Gostojić (2020)	Methodology (input data)	The paper identifies the relevant types of judicial datasets, reviews widely used OGD assessment methodologies, assesses the openness of these datasets in selected countries, and proposes corrective actions to improve effectiveness and efficiency of open data initiatives.	GODI, ODB
McBride et al. (2018)	Background / state of the art / Discussion & Methodology (input data)	The paper inspects Estonia's low maturity of the OGD in relation to a highly developed e-government, through an exploratory case study on Estonia, using document analysis, survey data and semi-structured interviews.	ODB, ODMR, OURdata Index
Musyaffa et al. (2018)	Background / state of the art / Discussion & Methodology (input data)	The paper presents the Open Fiscal Data Publication framework called OFDP for assessing the quality of open fiscal datasets based on the survey resulted main quality factors to be included in it. The authors identified the quality issues that are typical for these datasets.	GODI, ODB
Musyaffa et al. (2020)	Background / state of the art / Discussion	Comparative analysis of linked open fiscal data, including analysis of the use of specific ontology to ensure uniform representation of open fiscal data publicly available data to enrich the context of datasets and establish relationship links between similar concepts in different datasets.	GODI, ODB
Neumaier et al. (2016)	Background / state of the art / Discussion	Authors propose a generic metadata quality assessment framework for the various open data portals, where data portals are processed independently of the portal software frameworks. Several quality metrics have been defined that can be assessed automatically and	GODI, Open Data Monitor

		effectively.	
Nikiforova (2020)	Background / state of the art / Discussion	Analysis of the usability of the Latvian open data portal, which aims to find key challenges that can negatively impact user experience. The usability framework was applied to the portal, thus identifying the weakest aspects of the portal and providing corrective actions.	GODI, ODB, ODMR
Nikiforova and McBride (2021)	Background / state of the art / Discussion	Usability analysis of 41 OGD portals by applying for them a unified framework, involving 40 individual users. The authors have developed an initial comparative international ranking of OGD portal usability and identified commonly occurring portal usability strengths and weaknesses across contexts.	GODI, ODB, ODMR, OURdata Index
Nikolov and Krumova (2019)	Background / state of the art / Discussion & Methodology (input data)	Based on an overview and analysis of literature, the authors divide EU countries into three clusters – changers, observers, and moderators. Although for e-government the correlation with Hofstede model is high, for open government it is significantly lower.	GODI, ODB, ODMR, ODRA
Park and Oh (2019a)	Background / state of the art / Discussion	Authors explore key factors influencing the openness of government, develop a Global Government Openness Index (GGOI) to assess the progress of government openness, and explore how these factors contribute to the development and evolution of individual open governments and groups of country by income.	GODI, ODRA, WJP OGI
Park and Oh (2019b)	Background / state of the art / Discussion	Develops a Government Openness Index (GOI), explore the relationship of variables in the GOI, and assess the relationships between the GOI and income levels of developing countries.	GODI, WJP OGI
Ponce and Ponce Rodriguez (2020)	Methodology (input data)	Authors developed a cross-section analysis to examine institutional, political, and economic factors that identify cross-country differences in the provision of OGD.	GODI, ODB, ODIN, Open Data Impact
Quarati et al. (2021)	Background / state of the art / Discussion & Methodology (input data)	The paper provides an overview of the use of selected OGD portals. In addition, the authors assess the potential impact of metadata quality on the use of geospatial datasets by assessing a correlation of metadata quality and data usage.	GODI, ODB, OURdata Index
Saxena (2017)	Background / state of the art / Discussion	The paper reviews the strengths and weaknesses of OGD for the Gulf Cooperation Council member states, namely Bahrain, Kuwait, Qatar, Oman, Saudi Arabia, and the United Arab Emirates.	GODI, ODB, ODIN, ODMR, OURdata Index, Open Data Monitor
Sołtysik-Piorunkiewicz and Zdonek (2021)	Background / state of the art / Discussion	The paper identifies trends and keywords in promoting open data for their use in Industry 4.0 and Society 5.0. The authors identify leaders in Europe in promoting the use of the open data in the context of Industry 4.0 and Society 5.0.	GODI
Thorsby et al. (2017)	Background / state of the art / Discussion	The paper investigates the features and content of open data portals in American cities. They are divided into five pre-developed categories and describe these portals, namely the Open Data Portal Index (ODPI), Data Content Index (DCI), a compilation of the two (Overall Index), the Number of Datasets and Number of Datasets per 100,000.	GODI, PSI Scoreboard
Vancauwenberg he et al. (2018)	Background / state of the art / Discussion	The paper introduces an Open Spatial Data Infrastructure (SDI) assessment framework to assessing the openness of spatial data infrastructures.	GODI, ODB, ODRA, ODMR, Open Data 500, Open Data Impact Map
Veja et al. (2021)	Background / state of the art / Discussion	The paper inspects the educational datasets of 28 German Open Government Portals of the city-, state-, and national- level, analyzing their quality. It identifies key shortcomings of existing open data portals in terms of the quality of datasets and the lack of their findability and granularity.	GODI, ODB
Wang and Shepherd (2020)	Background / state of the art / Discussion (partly Methodology (input data))	The paper assesses the degree of openness of the OGD in the UK as a leader in the OGD movement, by examining a sample of 400 datasets available on the OGD portal.	GODI, ODB
Wang et al. (2018)	Background / state of the art / Discussion & Methodology (input data)	The paper develops a prioritized assessment framework for local Chinese OGD portals and validates its capabilities based on a case study.	GODI, ODB, OURdata Index, Open Data census, Open Data Monitor
Welle Donker and van Loenen, (2017)	Background / state of the art / Discussion	Authors developed a system / framework that assesses the supply of open data, open data governance, and the characteristics of open data user. It can be used to indicate on the aspects of the open data ecosystem that require attention and improvements.	GODI, ODB, ODMR, PSI Scoreboard
Wen and Hwang (2019)	Methodology (input data)	The paper reviews the levels of OGD in different countries compared to the levels of development of those countries. Identify those countries, which OGD level do not correspond to their development level.	OGDI (census)

Wiedenhoft et al. (2019)	Methodology (input data => new model)	The indices are used as a secondary data forming an indicator. The aim is to identify the contribution of OGD disclosure and debirocracy to reducing the level of corruption in the country. The authors have established a theoretical model with three global country-based indicators - the Corruption Perception Index, GODI, and The Ease of Doing Business Ranking.	GODI
Zhao and Fan (2021)	Background / state of the art / Discussion	A fuzzy-set qualitative comparative analysis to explore the mechanisms and their main aspects for interaction and coordination between different conditions on OGD performance.	GODI
Zheng et al. (2020)	Methodology & Background / state of the art / Discussion	Seven OGD assessment "programs" are used to systematically compare and analyze the frameworks, indices, and methods. The paper proposes a new index - the OGDI. The paper also covers a brief comparison of current international practices on evaluating OGD, i.e., indices and rankings.	GODI, ODB, ODRA, ODMR, OURdata Index, PSI Scoreboard
Žuffová (2020)	Methodology – input data	Empirical study on the relationship between OGD against FOI laws and corruption.	GODI, ODB
Zuiderwijk et al. (2021)	Methodology & Background / state of the art / Discussion	The paper compares metrics and methodologies used to measure, benchmark and rank governments' progress in OGD initiatives. Using a critical meta-analysis approach, the authors compared 9 benchmarks. The study concludes that both existing OGD benchmarks and academic open data progress models use very different metrics and methodologies, although the impact of open data is generally not measured.	GODI, ODB, ODIN, ODRA, ODMR, WJP OGI, Open Data Economy by Capgemini Consulting, OGD by The Economist Intelligence Unit

A further in-depth analysis of the most relevant papers included in the literature analysis revealed that in some cases a few more indices and rankings are being used. Therefore, here we cover some of them with their key points and a brief explanation of why they are not suitable for our study but can be used in other contexts.

The **ODRA**, launched by the World Bank's Open Government Data Working Group, is a methodological tool that helps to plan actions that the government authority could consider establishing an Open Data program at different administrative levels. It provides eight dimensions, including contexts and questions to ask, to successfully implement the Open Data program (World Bank Group, 2015). The ODRA is related to the GODI from 2013 from which it reuses the key types of datasets. It also proposes to consider other indices, such as the ICT Development Index published by the ITU. However, although this assessment tool can be useful for countries in their open data efforts, there is no output that would enable to benchmark and rank countries based on this framework, as well as no other update of the methodology from 2015. **PSI Scoreboard** is a tool to measure the status of open data and PSI reuse throughout the EU. Its goals are based on the EU's directives from 2003 and 2013 which provided a common legislative framework for this area. The latest Scoreboard is from 2013 and, among other, it assessed formats, pricing, or events and activities. Since 2015, the assessments around open data in the EU are directed towards the ODMR.

The **WJP OGI** was published only once in 2015 and it is conceptually derived from the WJP Rule of Law Index, which is published annually from 2008. The WJP OGI is focused on public experiences and perceptions, where four dimensions are important, i.e., publicized laws and government data, right to information, civic participation, and complaint mechanisms. Each of them is scored from 0 to 1 based on replies to household surveys and in-country expert questionnaires (World Justice Project, 2015). Although we, as many researchers, share their opinion that efforts can be measured more correctly when an end-user is involved, this index is out of scope of this study being very diverse with those that are more classic.

Open Data 500 launched in 2014 under GovLab focuses on 500 U.S. companies that use OGD as a key business resource. This tool has provided a way to map and visualize relationships between companies and open government datasets, and to carry out analysis of other sectors. However, to the best of our knowledge, this tool is no longer supported. Dawes et al. (2016) classifies it as a use- or user-

oriented approach. Another GovLab project is the **Open Data Impact** – a repository that provides a detailed understanding of the various processes and factors underlying the demand, supply, release, use and impact of open data, and assesses and provides evidence that open data have the potential to influence society in different favorable ways. In addition, they provide practical insights to policy makers, representatives of civil society, businesses, researchers and others who want to open or use open data, i.e., forming a handbook (GovLab, 2021).

The **Open Data Impact Map** is a public database that includes organizations using the OGD from around the world, launched by Open Data for Development Network (OD4D). The developers of this database stress that they do not attempt to rate, assess, or quantify the economic or social value of open data, nor do they provide a random or representative sample of use cases. They also do not guarantee the accuracy of all entries, as data are collected in three ways – (1) a web-based survey, (2) using a network of regional supporters representing 20 countries providing examples and insights based on local expertise, (3) researches. The **Open Data Census** is a platform operating by Open Knowledge International and closely associated with the GODI forming a so-called Open Data Survey. It is used to compare the progress of different cities and local areas in disseminating Open Data.

Open Data Monitor is a framework that retrieves datasets by collecting metadata from different open data sources. It is potentially able to analyze and visualize metadata and, as a result, to discover the hidden potential of existing resources by identifying gaps that require additional open data. However, the complete data are available only for 2015 and there is no mention about the scores and that the framework has been changed since then. Both the Open Data Census and the Open Data Monitor are classified as data-oriented approaches, focusing on the publishing policies of the countries they considered (Dawes et al., 2016). **EGovOI Index** is part of the OG benchmark model (OpenGovB) proposed by Veljković et al. (2014). It examines the openness of e-government looking at their state of governmental efforts, their readiness to publish open datasets in a timely manner, orientation toward users' needs and their involvement in government. Although the appears to be sufficiently detailed and useful to assess the efforts of each country, we have not covered it in this paper because it does not provide a national ranking, but rather it should be used by governments or enthusiasts to assess these efforts for their own countries while we are more interested in cross-country comparisons / rankings.

4.3 Summary of the findings

4.3.1 The structure

To clearly summarize the structure of open data indices and rankings resulting from the previous sections, we must first define the terms that relate to this structure and how it is presented towards governments and other stakeholders. Based on our findings, both index and rank are often used interchangeably but the term "index" is mostly in the title of the concrete index. The main purpose of the index is usually to rank countries or other entities. However, since the most indices are accompanied with a report, the term "benchmark or benchmarking" represents the goal of the index more precisely. To encompass all the elements and their relationships in the open data ecosystem, the indices are constructed as composite measures, i.e., that their benchmarking framework can be structured differently. We identified the following levels that explain this structure:

1. The highest level is represented by a score, i.e., a final score made by combining several sub-indices, dimensions, pillars, indicators etc.;

2. Dimension = sub-index = pillar etc. represents various number of levels and aspects on which the score is built;
3. Indicator = category = component represents various types of variables, composite and simple indicators;
4. The lowest level is characterized by a metric, i.e., a simple measure that is represented by a value for each entity.

Table 14 answers the RQ1 by providing a summary of the structure and weights of the indices through the years. Open data indices and rankings described in detail in this paper use a wide range of qualitative and quantitative methodologies and many different data collection techniques: sampling, surveys, crowdsourcing, expert surveys, peer-reviews, and analysis of the literature, among others. Based on data from Table 14, we can divide those indices into three groups. Those which measure: 1) openness of selected data categories (GODI, ODIN), 2) different aspects of open data ecosystem through a large number of variables (OURdata Index, ODMR), and 3) those which try to combine both approaches (ODB, OGDI). In general, variables used by the rankings describe the political and legal environment (data provision and reuse), national portal, quality of data and data provision systems, as well as environmental, social, political, and economic impact. This also makes it possible to determine the relevance of the index under question and facilitates more accurate index selection depending on the task posed.

Table 14: Summary of the structure and weights of the indices through the years.

Highest level	**Year**	**Dimensions**			**Lowest level**
		Name	Weight (original)	Weight (standardized)	Total number
GODI	2013 2014	technical openness legal openness	50 50	50% 50%	10 categories of data, 9 assessment criteria
	2015	technical openness legal openness	50 50	50% 50%	13 categories of data, 9 assessment criteria
	2016	technical openness legal openness	60 40	60% 40%	15 categories of data, 6 assessment criteria
ODB	2013	readiness implementation impacts	1/5 3/5 1/5	20% 60% 20%	18 variables, 15 categories of data, 10 assessment criteria
	2014	readiness implementation impacts	1/4 2/4 1/4	25% 50% 25%	18 variables, 15 categories of data, 10 assessment criteria
	2015 2016 2017	readiness implementation impacts	35% 35% 30%	35% 35% 30%	20 variables, 15 categories of data, 10 assessment criteria
OURdata Index	2015	data availability data accessibility pro-active support	33.3% 33.3% 33.3%	33.3% 33.3% 33.3%	19 variables
	2017 2019	data availability data accessibility pro-active support	33.3% 33.3% 33.3%	33.3% 33.3% 33.3%	170 variables
ODIN	2015 2016	social statistics economic statistics	33.3% 33.3%	33.3% 33.3%	20 categories of data, 10 assessment criteria

		environment statistics	33.3%	33.3%	
	2017 2018	social statistics economic statistics environment statistics	33.3% 33.3% 33.3%	33.3% 33.3% 33.3%	21 categories of data, 10 assessment criteria
	2020	social statistics economic statistics environment statistics	33.3% 33.3% 33.3%	33.3% 33.3% 33.3%	22 categories of data, 10 assessment criteria
ODMR	2015	policy impact portal	500 300 450	40% 24% 36%	24 variables
	2016	policy impact portal	530 300 510	39.55% 22.39% 38.06%	26 variables
	2017	policy impact portal	620 300 580	41.33% 20% 38.67%	26 variables
	2018	policy impact portal quality	27% 26% 26% 21%	27% 26% 26% 21%	20 variables
	2019 2020	policy impact portal quality	25% 25% 25% 25%	25% 25% 25% 25%	60 variables
OGDI	2020	policy and institutional framework national portal data availability	20% 25% 55%	20% 25% 55%	5 categories of data, 6 variables

Weight (original) – weight used by index / rank (column 4),

Weight (standardized) – weight, recalculated as a percentage (column 5)

The ODIN used the most coherent methodology over the years. Changes were limited to a slightly increasing number of data categories or switching them between dimensions. The scope of this Index can easily explain it: it is limited to statistical data provided by the national statistical offices, which means that it measures the openness of data that are very well structured and accompanied by standardized metadata, which makes the assessment much easier. The same cannot be said about the GODI, which went through several changes over the years, reducing the assessment criteria from nine to six and increasing the number of included categories from ten to fifteen. As a result, the increase of the scope was possible by simplification of the measurement method.

From all the rankings analyzed in this paper, the largest increase of the scope can be noticed in the OURdata Index, which jumped from 19 variables used in 2015 to 170 variables in 2017. To understand this change, we must refer to the political context. The Index, created by the OECD to measure open data ecosystems in its member states, was based on G8 Open Data Charter (OECD, 2015) in its first iteration and was considered a pilot study. The following reports were built around International Open Data Charter (ODC, 2015). The second Index in this group, the ODMR, offers the biggest methodological puzzle, with only data presented in the last two reports being comparable. The number of variables used by the ODMR changed over years and the scores / weights used in the

consecutive years. The reasoning behind those changes offered by the authors is very limited, and the adjustments of the scores seem to be sometimes erratic: e.g., subdimension "usability of the portal" is assigned a maximum of 100 points in 2015, down to 60 points in 2016, and up to 90 points in 2017. The reboot of the methodology in 2018 led to more changes in 2019; however, they seem to be adjustments introduced to make the report clearer and easier to read and do not change the general philosophy behind the Index.

Third group, which includes reports combining both approaches, more precisely the ODB and the OGDI. The ODB uses consistent methodology over the years. The changes include adjustments to weights assigned to dimensions and an increase of variables from eighteen to twenty in 2015. Most of the changes are limited and do not interfere with the comparability of data over the years. Unfortunately, only one edition of the OGDI has been published, making it impossible to analyze in-depth, especially since it does not provide a detailed description of the methodology used.

Apart from different approaches and methodologies used, it is also important to underline that analyzed reports cover different periods. As a result, 2015 was covered by the highest number of reports – five, while the most recent 2020 is reflected only in three reports.

4.3.2 The development

As regards the RQ3, the beginning of benchmarking open data efforts has its origin in the concept of PSI, more precisely in 2003 when the PSI Directive was released, and the follow up PSI Scoreboard. Although the requirements on open data were still not clearly defined and accepted in this period, the concept of PSI has introduced the conditions to reuse data disclosed online by public sector agencies and institutions. A set of eight principles of OGD was published in 2007. In 2010, Tim Berners-Lee established a 5-star scheme for measuring the quality of open data on the web. An Open Government Partnership was launched in 2011 to promote openness, transparency, accountability, and the use of modern ICT in governance. All these requirements quickly proliferated into a variety of areas and topics, implemented by both private and public sectors, and enabled to form benchmarking frameworks that began to be widely accepted by governments worldwide.

The biggest rise of open data indices and rankings can be found between 2013 and 2015 when most of the publishers aimed to cover the requirements on open datasets in their frameworks. The number and focus of datasets categories covered, respectively their importance (weights) based on the demand from users, has changed over the years. It is characteristic for this period that all sources of OGD and open datasets on the web were considered. This sometimes resulted in situations where some countries got better scores since their sources were easily findable and their websites more usable. However, this could affect the real state and development of involved countries. On the other hand, a lot of countries began to launch open data portals where OGD from national, regional, or local levels were centralized and thus governments could focus more on the quality of open datasets and features of portals needed to work with them.

After 2015, open data portals have become the center of benchmarking frameworks and their use by users was considered as a key part of measures dealing with impact and value. Some changes are also related to the inclusion of an OGD impact dimension in cases where it had not been previously considered or increasing its weight. This could be because the impact is becoming more significant, since majority countries have already established at least some basic elements of the OGD policies, established the open data portals and advanced them to their full operating, and now feel that there is time when the impact of the OGD and open data portals should be measured, i.e., whether it is and what is the level of their impact. This assumption is also in line with an overview we have provided,

including the ODB, OURdata Index, etc., and GODI which have eliminated dimensions such as the existence of data and their availability in digital form, which are now considered to be fulfilled by default, with a greater focus on more technical aspects closely related to open data and the principles of openness closely related to quality and further reuse of data with their subsequent respective impact.

We can also identify several other dimensions of open data ecosystem that were incorporated into the frameworks. These deal with the processes that support participation and collaboration of users to create value. Most of them are provided by open data portals or other channels and platforms used by public sector agencies and institutions. However, although the importance of what users want and need is usually the first criterion, there are also other stakeholders that influence the ecosystem. Governments actions and policies are crucial to orchestrate it and formalize the existing processes behind OGD. However, the preference of one of these groups is affected by the goal of the index or rank and thus we can find different dimensions that are used by different frameworks.

In the last three or four years, environmental issues and sustainability have the biggest impact on what is measured and how the benchmarking frameworks are updated. It is closely related to the problematics of data infrastructure consolidation, centralization of resources, and green computing. Although open data are often seen only as a service or resource that should be freely available and easily findable by users while meeting appropriate standards, there is always an infrastructure behind them that comprises hardware, software, as well as human resources. All these elements must be considered while assessing the open data ecosystem and its efficiency in gaining potential value and a quantifiable impact on society. We therefore argue that these rankings, which assess all the above-mentioned aspects, could be considered more universal, although they still cannot be considered to be one-fits-all, because acquiring of a very comprehensive view on the current state is a multi-dimensional and multi-perspective task that could be obtained mainly in the context of one country, where different types of users and different types of rankings are involved. Therefore, the assessment of the efforts of one individual country is likely to be more accurate if the ranking is combined with a detailed country-specific analysis, see Nikiforova and Lněnička (2021) for an example, thereby ensuring better granularity.

To sum up, a new era of open data may be considered to have begun, when the paradigm has changed significantly, and the questions previously considered to be debatable became daily reality and even routine. Finally, the need to update the framework is usually solved by increasing or decreasing the weight of a dimension or indicator or by removing one of them and / or replacing it by another. It can be concluded that the changes that have occurred over years in benchmarking frameworks help improve scores countries that were a step forward in their open data efforts. This is especially the case of open data portals and their quality including usability and the range of features. Furthermore, it is the sustainable development which affects the impact and efficiency dimensions of open data indices and rankings in recent years.

5 Discussion and limitations

Discussing the development is always difficult since each period is influenced by different factors. In our paper, we identified four main periods that characterize how open data efforts are benchmarked. Although there can be overlaps between the periods, they provide us with an overview of *what were and still are the priorities, what had to be changed and updated regarding new trends in ICT as well as behavior and needs of users and other stakeholders, what are the best practices,* and *what way the efforts to benchmark open data will be directed in the future*.

The first period, before the concept of open (government) data was clearly stated, is represented by the PSI Scoreboard in our paper. However, there are also other indices and rankings that could affect the development of open data. Some of them were introduced in the 90s and are still active, at least in relation to the areas they deal with. They focus mostly on transparency and democracy metrics that are tied to information and Internet society and how information flows and corresponding processes are addressed. In addition, e-government indices and rankings have their merits on shaping the way the benchmarking open data efforts has taken. But this period is very difficult to be described clearly since there were many separate and unique factors and pressures that formed these beginnings.

The following two periods are characterized by emerging sets of principles and best practices around open data and open government. These were introduced by different organizations and countries can join them and participate on improvements in corresponding areas. Accordingly, countries that joined one of these partnerships received points to the overall score. However, nowadays, they lose their significance and are seen only as platforms for discussions and debate. The attention is directed towards data infrastructures in which open data portals are the central point for working with open datasets. If the infrastructure is built and the portal provides appropriate features in a usable and useful manner, governments and data publishers focus on promoting their efforts in various initiatives, hackathons, and courses that should improve data literacy and skills of users.

The structure and the actors to be involved in the process of benchmarking is other issue. In some cases, there is a trend to involve end-users. This is in line with Wang and Shepherd (2020), according to which existing frameworks, indices, and rankings are inadequate because they assess different aspects of the openness of the OGD, such as service and portal, their characteristics and quality, rather than the openness of datasets from the perspective of the ordinary citizen. On the other hand, it can be argued about what categories of datasets are high-value for users and should be included in the frameworks. Also, features provided by open data portals may not correspond to the requirements and needs of users but also their skills. There are various studies that deal with these issues, such as de Juana-Espinosa and Luján-Mora (2019), Lněnička and Nikiforova (2021), Nikiforova and Lněnička (2021) or Nikiforova and McBride (2021),

The last period that currently affects the benchmarking frameworks most should be discussed into the context of future changes and trends. It is especially the pressure on the impacts of open data reuse in various areas. Efficiency and effectiveness of resources and sustainability and smartness of solutions are goals that surpassed transparency, accountability, and engagement of stakeholders. More attention is focused on these impacts of open data and their reuses. It is not only about the number of services and applications based on open datasets but how they support those new goals. In this regard, we suppose that these trends will be incorporated into benchmarking frameworks, including using data from sensors and Internet of Things, while using Artificial Intelligence and Big Data analytics to transform these data to meet appropriate standards and be available on open data portals in real-time thereby supporting and facilitating the development of the Industry 4.0 and Society 5.0, also in line with Sołtysik-Piorunkiewicz and Zdonek (2021) and Nikiforova (2021).

The changes in methodologies are to be expected if the measurement is conducted over a longer period, especially in the open data ecosystem, defined and redefined by updated political goals and technological innovations. Based on our findings, we can state that the structure and / or the methodology behind indices and rankings that published at least two editions has changed. Generally, if the topic is new, it usually evolves quickly based on best practices and more detailed specifications on how to measure it. In addition, since the topic of open data heavily relies on ICT, it must deal with the pressures represented by modern trends this area and the way how users consume information and

work with data. However, those adjustments should be introduced with a detailed explanation, which is not always the case and is one of the most significant limitations of this paper.

In other words, in some cases, differences in results are not the subjects of changes just in open data efforts, since in many cases they are changing (mainly reduces) because of methodological changes, which usually follow the latest trends, as new requirements are usually implemented relatively quickly by developed countries with serious gap for other countries. In some cases, methodologies are subject to changes not because other aspects need to be addressed, but to make them more readable and easier to interpret as was the case for ODMR. Even more, in some cases, benchmarks represent a combination of external indices, and their results are even more difficult to understand and interpret (see OGDI). As regards average results for specific indices, the latest releases cover a higher number of countries where additional countries typically represent developing countries, which have a negative impact on the average result, since initial releases are often focused mainly on developed countries. This was the case for GODI, for instance.

In some cases, mainly when the openness of (statistical) data is considered (see section on ODIN), it may be more likely that results will be compared over the years, and these average scores tend to increase, i.e., results improve annually and at the same time contribute to the overall growth of the open data ecosystem. This makes it difficult to use benchmarks when it is necessary to track the changes over time, since it is sometimes very difficult to distinguish between them when the results in the scope of specific index are or are not comparable.

Our overview of the existing rankings points to differences in the coverage of these rankings. i.e., although today international cooperation and the exchange of knowledge and experience between different countries and continents are becoming popular, some of them limit this cooperation to the region. This could have a negative impact on the overall development of the open data ecosystem, as it makes it a challenge to find cooperation points, which could potentially be done through these rankings, i.e., identifying key areas to focus on etc. Moreover, studies focusing on the same area are more likely to be carried out several times, having different samples, and trying to make their findings more generalizable, although in a more decentralized way. However, it should be noted that some rankings extend the list of countries to be covered from one edition to another.

Most of the indices and rankings analyzed are published every year or once in two years. It can be argued about what time frame is the best and if it can really cover the changes in the open data ecosystem, if there are any. This may be also the reason why some of them are on a hiatus or probably will not be published again. On the other hand, if the organization behind the index or rank fulfill its goals, then, it is not necessary to publish next report.

This issue has also caught the attention of other researchers, i.e., there are several studies aimed to explore the topic of benchmarking open data efforts. Zheng et al. (2020) examined seven OGD benchmarks, systematically comparing and analyzing relevant frameworks, indices, and methods. Based on their results they produced the OGDI. Although they analyzed the structure they did not deal with the development. Zuiderwijk et al. (2021) explored nine OGD benchmarks using a critical meta-analysis approach. Wang et al. (2018) used selected dimensions of open data indices and rankings to propose a framework for assessing local-level OGD portals. Kawashita et al. (2020) focused on different dimensions that OGD benchmarks assess to find out what aspects they measure to help practitioners and researchers decide which benchmark is best suited for a specific purpose.

As regards the limitations, the most important one is the interpretation of the results. As reported by Bannister (2007), "*the outcomes of benchmarks need to be interpreted sensibly and it is always*

necessary to be aware of the risks of their politicization." This leads to a situation where governments focus their open data initiatives and data portals to maximize points on open data indices and rankings (Nikiforova and McBride, 2021). But sometimes organizations publishing the indices and rankings are reluctant to change the framework or methodology to enable comparisons over time. When the publishers are reluctant to change their frameworks based on the feedback, new ones are introduced by researches or national governments. It is also because existing indices and rankings are sometimes too universal and focus on a set of concepts without distinction between them. As a result, the value of these outputs is reducing because it is difficult to understand what is the weakest aspect to be improved when this set composed of a number of different aspects.

A review of literature carried out in this study revealed that, despite our assumption about the scope of the use of indices, according to which they mainly serve as support for the discussion of authors or the motivation for studies, they are often used as an input or to produce new methodologies. These rankings are found to be important not only for politicians, OGD portal holders, OGD policy developers etc., but also for researchers and enthusiasts, by closely monitoring the state of the open data and respective portals as well as trying to improve them. Unfortunately, sometimes they fail in this respect, because existing indices are rarely thoroughly explained in terms of both the overall methodology, including the way in which data for the index were collected and the specific calculations that we have also experienced during the study. This is even worse when the methodology has changed, but politicians, researchers, enthusiasts were unaware of it and treat the results just as they had previously, which negatively affect decision-making by providing inaccurate results and conclusions with consequent plans and actions. However, this study, will hopefully be supportive material for such studies, as we have managed to reveal and summarize some of these data.

From the summarization of the development we made, it is obvious that the first step how to face new trends in ICT and other challenges is to increase or decrease the weight of a single metric on the lowest level of the framework, then change the weights of indicators and / or dimensions. If this is not sufficient, then, new metrics / indicators / dimensions are added, or the old ones are removed. To sum this up, the key enablers of changes are the principles and requirements on open (government) data, open data portals, efforts for a collaborative and participatory environment, and sustainable development.

6 Conclusions

Although each country has its own specifics and should focus on the needs of its citizens and businesses, the globalization process and the need for transnational cooperation create ecosystems where the corresponding processes need to be standardized. These pressures are then reflected in both the private and public sectors. The public sector sets up rules for the environments and ecosystems, in which it then cooperates with citizens and businesses. It is therefore the main player who should assess the impact of its policies and strategies. In this context, identification of key areas as well as successes and failures are necessary. Thus, benchmarking is increasingly important for the development of individual countries.

This study aimed to explore the approaches to benchmark and measure open data efforts on the level of countries. The conducted research aimed to answer three primary RQs. First, we have identified six ranking systems to be popular and widely used – GODI, ODB, OURdata Index, ODIN, ODMR, and OGDI, which were further thoroughly inspected by analyzing the methodologies and indicators lying under them and how they have changed over time. This allows to conclude whether the results of the different releases of the same index can be comparable and used as a basis for decision-making

on the development of specific aspects and input data for the definition of further activities on OGD. Indices and rankings discussed in this paper use two approaches: measuring the openness of selected data categories and using variables to describe the broader context of the open data ecosystem although the combination of them both is allowed. The first approach is data-centered and is limited to the provision side. It uses mostly binary variables and could be fully automated soon. The second approach offers a more complete picture of the open data ecosystem in each country, measuring data supply, usage, and impact.

As regards the second RQ, a systematic literature review was carried out on five leading digital libraries. It led us to the conclusion that although the indices are rather supportive material in articles that allow authors to justify their findings or stress the importance of the study, primary studies often use existing rankings and their results as an input to their methodologies and sometimes combine different indices, producing new ones. Finally, as far as the third RQ, we found that vast majority of indices aimed at assessing progress over time are continuously updating their methodologies to meet current requirements / prerequisites that makes current results more relevant to current ICT trends. There we expect the closest connection of new releases with such trends as sensor-generated data, IoT, Big data, and their closer link to Society 5.0, as well as the increased popularity of their use in the context of AI. However, these changes also make the results of different releases of the same index incomparable.

The output of this study is the classification of selected benchmarking frameworks, analysis of their development, components, and description. An overview of their use in practice, critical comments, limitations, and possible modifications that improve them are also included. Our further research builds on and explores these findings using a quantitative approach to introduce a new framework and its validation on selected countries.

References

Aarshi, S., Malik, B. H., Habib, F., Ashfaq, K., Saleem, I., & Tariq, U. (2018). Dimensions of open government data web portals: a case of Asian countries. *International Journal of Advanced Computer Science and Applications*, 9(6), 459-469. https://doi.org/10.14569/IJACSA.2018.090663

Ahonen, P. (2018). Learning from opening data in the context of e-governance: Finland, with special reference to government location data. In: Alcaide Muñoz L., Rodríguez Bolívar M. (Eds.) *International E-Government Development* (pp. 271-290). Palgrave Macmillan, Cham. https://doi.org/10.1007/978-3-319-63284-1_11

Alarabiat, A., Soares, D., Ferreira, L., & de Sá-Soares, F. (2018). Analyzing e-governance assessment initiatives: an exploratory study. In *Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age* (pp. 1-10). https://doi.org/10.1145/3209281.3209309

AlRushaid, M. W., & Saudagar, A. K. J. (2016). Measuring the data openness for the open data in Saudi Arabia e-Government: a case study. *International Journal of Advanced Computer Science and Applications*, 7(12), 113-122. https://doi.org/10.14569/IJACSA.2016.071215

Altayar, M. S. (2018). Motivations for open data adoption: an institutional theory perspective. *Government Information Quarterly*, 35(4), 633-643. https://doi.org/10.1016/j.giq.2018.09.006

Amalia, D. N., & Susanto, T. D. (2019). Analysis of motivation and perceived risk factors in open data measurement: a conceptual model. *Procedia Computer Science*, 161, 308-315. https://doi.org/10.1016/j.procs.2019.11.128

Aquaro, V. (2020). *E-Government Survey 2020 Digital Government in the Decade of Action for Sustainable Development*. New York: Department of Economic and Social Affairs, United Nations.

Aquaro, V., Barthélemy, M., & Schweinfest, S. (2018). *United Nations e-Government Survey 2018*. In United Nations, Department of Economic and Social Affairs. New York: United Nations.

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, 32(4), 399-418. https://doi.org/10.1016/j.giq.2015.07.006

Bannister, F. (2007). The curse of the benchmark: an assessment of the validity and value of e-government comparisons. *International Review of Administrative Sciences*, 73(2), 171-188. https://doi.org/10.1177/0020852307077959

Blank, M. (2019a). *Measuring open data maturity*. Retrieved from https://www.europeandataportal.eu/sites/default/files/method-paper_insights-report_n5_2019.pdf

Blank, M. (2019b). *Open Data Maturity: Report 2019*. Retrieved from https://doi.org/10.2830/073835

Bonina, C., & Eaton, B. (2020). Cultivating open government data platform ecosystems through governance: Lessons from Buenos Aires, Mexico City and Montevideo. *Government Information Quarterly*, 37(3), 101479. https://doi.org/10.1016/j.giq.2020.101479

Carrara, W., Fischer, S., & Steenbergen, E. (2015). *Open Data Maturity in Europe 2015: Insights into the European state of play*. https://doi.org/10.2830/23574

Carrara, W., Nieuwenhuis, M., & Vollers, H. (2016). *Open Data Maturity in Europe 2016: Insights into the European state of play*. https://doi.org/10.2830/41983

Carrara, W., Radu, C., & Vollers, H. (2017). *Open Data Maturity in Europe 2017: Open Data for a European Data Economy*. https://doi.org/10.2830/918627

Corrêa, A. S., de Paula, E. C., Correa, P. L. P., & da Silva, F. S. C. (2017). Transparency and open government data: a wide national assessment of data openness in Brazilian local governments. *Transforming Government: People, Process and Policy*, 11(1), 58-78. https://doi.org/10.1108/TG-12-2015-0052

Crowell, J. (2016). *Open Data Inventory 2016: Methods Report*. Retrieved from https://odin.opendatawatch.com/Downloads/otherFiles/ODIN-2016-Methodology.pdf

Crowell, J. (2017). *Open Data Inventory 2017: Methodology Report*. Retrieved from https://odin.opendatawatch.com/Downloads/otherFiles/ODIN-2017-Methodology.pdf

Crowell, J. (2018). *Open Data Inventory 2018/19: Methodology Report*. Retrieved from https://odin.opendatawatch.com/Downloads/otherFiles/ODIN-2018-Methodology.pdf

Crowell, J., & Swanson, E. (2020a). *Open Data Inventory 2020/21: Executive Summary*. Retrieved from https://odin.opendatawatch.com/Downloads/otherFiles/ODIN-2020-ExecutiveSummary.pdf

Crowell, J., & Swanson, E. (2020b). *Open Data Inventory 2020/21: Methodology Guide*. Retrieved from https://docs.google.com/document/d/1MBK0hN6MoQrii7_E1bmRXmsUcE8Fbb-Q32nxm8d8qTw/edit

G8. (2013). *G8 Open Data Charter*. Lough Erne: G8.

OECD. (2015). *Government at a Glance 2015*. https://doi.org/http://dx.doi.org/10.1787/gov_glance-2015-en

Dawes, S. S., Vidiasova, L., & Parkhimovich, O. (2016). Planning and designing open government data programs: an ecosystem approach. *Government Information Quarterly*, 33(1), 15-27. https://doi.org/10.1016/j.giq.2016.01.003

de Juana-Espinosa, S., & Luján-Mora, S. (2019). Open government data portals in the European Union: considerations, development, and expectations. *Technological Forecasting and Social Change*, 149, 119769. https://doi.org/10.1016/j.techfore.2019.119769

Escobar, P., Candela, G., Trujillo, J., Marco-Such, M., & Peral, J. (2020). Adding value to linked open data using a multidimensional model approach based on the RDF data cube vocabulary. *Computer Standards & Interfaces*, 68, 103378. https://doi.org/10.1016/j.csi.2019.103378

Galeone, C., & Bonzi, R. (2020). Open data: applications in social care and health. *Statistica Applicata-Italian Journal of Applied Statistics*, 32(2), 125-139. https://doi.org/10.26398/IJAS.0032-008

Gonzálvez-Gallego, N., & Nieto-Torrejón, L. (2021). Government data openness and coverage. How do they affect trust in European countries?. *Journal of Data and Information Science*, 6(1), 139-153. https://doi.org/10.2478/jdis-2021-0010

GovLab. (2021). Retrieved from https://odimpact.org/about.html

Grzenda, M., & Legierski, J. (2021). Towards increased understanding of open data use for software development. *Information Systems Frontiers*, 23, 495-513. https://doi.org/10.1007/s10796-019-09954-6

Jetzek, T. (2016). Managing complexity across multiple dimensions of liquid open data: the case of the Danish basic data program. *Government Information Quarterly*, 33(1), 89-104. https://doi.org/10.1016/j.giq.2015.11.003

Kawashita, I., Baptista, A. A., & Soares, D. S. (2020). An assessment of open government data benchmark instruments. In *EGOV-CeDEM-ePart-* 2020: Linköping, Sweden* (pp. 29-40). CEUR Workshop Proceedings. Retrieved from http://ceur-ws.org/Vol-2797/paper2.pdf

Kim, H. (2019). Analysis of standard vocabulary use of the open government data: the case of the public data portal of Korea. *Quality & Quantity*, 53(3), 1611-1622. https://doi.org/10.1007/s11135-018-0829-z

Knippenberg, L. van. (2020a). *Measuring open data maturity, sixth edition 2020*. Retrieved from https://www.europeandataportal.eu/sites/default/files/method-paper_insights-report_n6_2020.pdf

Knippenberg, L. van. (2020b). *Open Data Maturity: Report 2020*. https://doi.org/10.2830/619187

Krotova, A., Mertens, A., & Scheufen, M. (2020). Open data and data sharing: an economic analysis *IW-Policy Paper*, 21/2020, Institut der deutschen Wirtschaft (IW), Köln.

Kubler, S., Robert, J., Neumaier, S., Umbrich, J., & Le Traon, Y. (2018). Comparison of metadata quality in open data portals using the analytic hierarchy process. *Government Information Quarterly*, 35(1), 13-29. https://doi.org/10.1016/j.giq.2017.11.003

Lafortune, G., & Ubaldi, B. (2018). *OECD 2017 OURdata Index: Methodology and results*. https://doi.org/https://doi.org/10.1787/2807d3c8-en

Lämmerhirt, D. & Brandusescu, A. (2019). Measurement. In T. Davies, S. Walker, M. Rubinstein, & F. Perini (Eds.), *The State of Open Data: Histories and Horizons* (pp. 320-338). Cape Town and Ottawa: African Minds and International Development Research Centre.

Lněnička, M., & Komárková, J. (2019). Big and open linked data analytics ecosystem: Theoretical background and essential elements. *Government Information Quarterly*, 36(1), 129-144. https://doi.org/10.1016/j.giq.2018.11.004

Lněnička, M., & Nikiforova, A. (2021). Transparency-by-design: What is the role of open data portals?. *Telematics and Informatics*, 61, 101605. https://doi.org/10.1016/j.tele.2021.101605

Luterek, M. (2020). Why are rankings of 'smart cities' lacking? An analysis of two decades of e-government benchmarking. In M. Themistocleous et al. (Eds.): *EMCIS 2020* (pp. 238-255). Springer, Cham. https://doi.org/10.1007/978-3-030-63396-7_16

Máchová, R. (2017). Measuring the effects of open data on the level of corruption. In Špalková, D., Matějová, L. (Eds), *Proceedings of the 21st International Conference Current Trends in Public Sector Research* (pp. 58-66). Masaryk University, Brno.

Máchová, R., & Lněnička, M. (2017). Evaluating the quality of open data portals on the national level. *Journal of Theoretical and Applied Electronic Commerce Research*, 12(1), 21-41. http://doi.org/10.4067/S0718-18762017000100003

Máchová, R., & Lněnička, M. (2015). Reframing e-government development indices with respect to new trends in ICT. *Review of Economic Perspectives*, 15(4), 383-411. http://doi.org/10.1515/revecp-2015-0027

McBride, K., Toots, M., Kalvet, T., & Krimmer, R. (2018). Leader in e-government, laggard in open data: exploring the case of Estonia. *Revue francaise d'administration publique*, 167, 613-625. https://doi.org/10.3917/rfap.167.0613

Marković, M., & Gostojić, S. (2020). Open judicial data: a comparative analysis. *Social Science Computer Review*, 38(3), 295-314. https://doi.org/10.1177/0894439318770744

Martin, S., Turki, S., & Renault, S. (2017). Open data ecosystems: Introducing the stimulator function. In A. Kő, & E. Francesconi (Eds.). *International Conference on Electronic Government and the Information Systems Perspective, EGOVIS 2017* (pp. 49–63). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-64248-2_5

Musyaffa, F. A., Engels, C., Vidal, M. E., Orlandi, F., & Auer, S. (2018). Experience: open fiscal datasets, common issues, and recommendations. *Journal of Data and Information Quality*, 9(4), 1-10. https://doi.org/10.1145/3190576

Musyaffa, F. A., Lehmann, J., & Jabeen, H. (2020). Cross-administration comparative analysis of open fiscal data. In *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance* (pp. 459-468). ACM. https://doi.org/10.1145/3428502.3428573

Neumaier, S., Umbrich, J., & Polleres, A. (2016). Automated quality assessment of metadata across open data portals. *Journal of Data and Information Quality*, 8(1), 1-29. https://doi.org/10.1145/2964909

Nikiforova, A. (2020). Assessment of the usability of Latvia's open data portal or how close are we to gaining benefits from open data. In *International Conferences Interfaces and Human Computer Interaction 2020* (pp. 51-58).

Nikiforova, A. (2021). Smarter open government data for society 5.0: Are your open data smart enough?. *Sensors*, 21(15), 5204. https://doi.org/10.3390/s21155204

Nikiforova, A., & Lněnička, M. (2021). A multi-perspective knowledge-driven approach for analysis of the demand side of the Open Government Data portal. *Government Information Quarterly*, 101622. https://doi.org/10.1016/j.giq.2021.101622

Nikiforova, A., & McBride, K. (2021). Open government data portal usability: a user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, 58, 101539. https://doi.org/10.1016/j.tele.2020.101539

Nikolov, H. S., & Krumova, M. Y. (2019). Hofstedes model in the context of e-government and open government in EU countries: countries clustering based on similarities and differences. *Smart Cities and Regional Development (SCRD) Journal*, 3(1), 29-46.

ODC. (2015). *International Open Data Charter*. Retrieved from https://opendatacharter.net/wp-content/uploads/2015/10/opendatacharter-charter_F.pdf

ODW. (2015). *The Open Data Inventory: Methodology*. Retrieved from https://doi.org/https://odin.opendatawatch.com/Downloads/otherFiles/ODIN-2015-Methodology.pdf

OKF. (2014). *Global Open Index: Methodology*. Retrieved from http://2014.index.okfn.org/methodology/

OKF. (2015). *Global Open Data Index: Methodology*. Retrieved from http://2015.index.okfn.org/methodology/

OKF. (2017). *Global Open Data Index: Methodology*. Retrieved from https://index.okfn.org/methodology/

Park, E. G., & Oh, W. (2019a). Factors and their relationships in measuring the progress of open government. *Aslib Journal of Information Management*, 72(1), 17-33. https://doi.org/10.1108/AJIM-04-2019-0083

Park, E. G., & Oh, W. (2019b). Developing a government openness index: the case of developing countries. *Information Development*, 35(1), 121-134. https://doi.org/10.1177/0266666917731946

Pérez, J. A. R., & Emilsson, C. (2020). *OECD Open, Useful and Re-usable data (OURdata) Index: 2019*. Retrieved from https://www.oecd.org/governance/digital-government/ourdata-index-policy-paper-2020.pdf

Ponce, A., & Ponce Rodriguez, R. A. (2020). An analysis of the supply of open government data. *Future Internet*, 12(11), 186. https://doi.org/10.3390/fi12110186

Quarati, A., De Martino, M., & Rosim, S. (2021). Geospatial open data usage and metadata quality. *ISPRS International Journal of Geo-Information*, 10(1), 30. https://doi.org/10.3390/ijgi10010030

Radu, C., & Cecconi, G. (2018a). *Landscaping Method - Overview*. Retrieved from https://www.europeandataportal.eu/sites/default/files/method-paper_insights-report_n4_2018.pdf

Radu, C., & Cecconi, G. (2018b). *Open Data Maturity in Europe: Report 2018*. https://doi.org/10.2830/593483

Rorissa, A., Demissi, D., & Pardo, T. (2011). Benchmarking e-government: a comparison of frameworks for computing e-government index and ranking. *Government Information Quarterly*, 28(3), 354-362. https://doi.org/10.1016/j.giq.2010.09.006

Saxena, S. (2017). Significance of open government data in the GCC countries. *Digital Policy, Regulation and Governance*, 19(3), 251-263. https://doi.org/10.1108/DPRG-02-2017-0005

Skargren, F. (2020). What is the point of benchmarking e-government? An integrative and critical literature review on the phenomenon of benchmarking e-government. *Information Polity*, 25(1), 67-89. https://doi.org/10.3233/IP-190131

Sołtysik-Piorunkiewicz, A., & Zdonek, I. (2021). How society 5.0 and industry 4.0 ideas shape the open data performance expectancy. *Sustainability*, 13(2), 917. https://doi.org/10.3390/su13020917

Susha, I., Zuiderwijk, A., Janssen, M., & Grönlund, Å. (2015). Benchmarks for evaluating the progress of open data adoption: Usage, limitations, and lessons learned. *Social Science Computer Review*, 33(5), 613-630. https://doi.org/10.1177/0894439314560852

Thorsby, J., Stowers, G. N., Wolslegel, K., & Tumbuan, E. (2017). Understanding the content and features of open data portals in American cities. *Government Information Quarterly*, 34(1), 53-61. https://doi.org/10.1016/j.giq.2016.07.001

Vancauwenberghe, G., Valečkaitė, K., Van Loenen, B., & Donker, F. W. (2018). Assessing the openness of spatial data infrastructures (SDI): towards a map of open SDI. *International Journal of Spatial Data Infrastructures Research*, 13, 88-100.

Veja, C., Hocker, J., Schindler, C., & Rittberger, M. (2021). Educational open government data in Germany: the landscape, status, and quality. In T. Schmidt, C. Wolff (Eds.): *Information between Data and Knowledge. Information Science and its Neighbors from Data Science to Digital Humanities* (pp. 370-396). Glückstadt, Verlag Werner Hülsbusch. https://doi.org/10.5283/epub.44955

Veljković, N., Bogdanović-Dinić, S., & Stoimenov, L. (2014). Benchmarking open government: an open data perspective. *Government Information Quarterly*, 31(2), 278-290. https://doi.org/10.1016/j.giq.2013.10.011

W3F. (2013). *Open Data Barometer. 2013 Global Report*. Retrieved from http://opendatabarometer.org/doc/1stEdition/Open-Data-Barometer-2013-Global-Report.pdf

W3F. (2015a). Open Data Barometer. 2014 Global Report. Retrieved from http://opendatabarometer.org/assets/downloads/Open Data Barometer - Global Report - 2nd Edition - PRINT.pdf

W3F. (2015b). *Open Data Barometer. ODB Methodology - v1.0 28th April 2015*. Retrieved from http://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-Methodology.pdf

W3F. (2016a). *ODB Global Report. Third edition*. Retrieved from https://opendatabarometer.org/doc/3rdEdition/ODB-3rdEdition-GlobalReport.pdf

W3F. (2016b). *Open Data Barometer: ODB Methodology - v1.0 28th April 2016*. Retrieved from https://opendatabarometer.org/doc/4thEdition/ODB-4thEdition-GlobalReport.pdf

W3F. (2018). *Open Data Barometer. Leaders Edition: From Promise to Progress*. Retrieved from https://opendatabarometer.org/doc/leadersEdition/ODB-leadersEdition-Report.pdf

Wang, D., Chen, C., & Richards, D. (2018). A prioritization-based analysis of local open government data portals: a case study of Chinese province-level governments. *Government Information Quarterly*, 35(4), 644-656. https://doi.org/10.1016/j.giq.2018.10.006

Wang, V., & Shepherd, D. (2020). Exploring the extent of openness of open government data–a critique of open government datasets in the UK. *Government Information Quarterly*, 37(1), 101405. https://doi.org/10.1016/j.giq.2019.101405

Welle Donker, F., & van Loenen, B. (2017). How to assess the success of the open data ecosystem?. *International Journal of Digital Earth*, 10(3), 284-306. http://doi.org/10.1080/17538947.2016.1224938

Wen, Y. F., & Hwang, Y. T. (2019). The associativity evaluation between open data and country characteristics. *The Electronic Library*, 37(2), 337-364. https://doi.org/10.1108/EL-04-2018-0081

Wiedenhoft, G., M. Luciano, E., & A. Magnagnagno, O. (2019). Data disclosure or debureaucratization: which is the best strategy to reduce corruption vulnerability?. In *Proceedings of the 20th Annual International Conference on Digital Government Research* (pp. 70-80). ACM. https://doi.org/10.1145/3325112.3325239

World Bank Group. (2015). *Open Data Readiness Assessment-Part B: Methodology*. Retrieved from http://opendatatoolkit.worldbank.org/docs/odra/odra_v3.1_methodology-en.pdf

World Justice Project. (2015). *The WJP Open Government Index Methodology*. Retrieved from https://worldjusticeproject.org/sites/default/files/ogi_methodology_0.pdf

Zhao, Y., & Fan, B. (2021). Understanding the key factors and configurational paths of the open government data performance: based on fuzzy-set qualitative comparative analysis. *Government Information Quarterly*, 38(3), 101580. https://doi.org/10.1016/j.giq.2021.101580

Zheng, L., Kwok, W. M., Aquaro, V., Qi, X., & Lyu, W. (2020). Evaluating global open government data: methods and status. In *Proceedings of the 13th International Conference on Theory and Practice of Electronic Governance* (pp. 381-391). https://doi.org/10.1145/3428502.3428553

Zhu, J., & Barthélemy, M. (2016). *United Nations e-Government Survey 2016*. In United Nations, Department of Economic and Social Affairs. New York: United Nations.

Zuiderwijk, A., Pirannejad, A., & Susha, I. (2021). Comparing open data benchmarks: which metrics and methodologies determine countries' positions in the ranking lists?. *Telematics and Informatics*, 62, 101634. https://doi.org/10.1016/j.tele.2021.101634

Žuffová, M. (2020). Do FOI laws and open government data deliver as anti-corruption policies? Evidence from a cross-country study. *Government Information Quarterly*, 37(3), 101480. https://doi.org/10.1016/j.giq.2020.101480