Open Urban Data and the Sustainable Development Goals

A small-scale Case Study on German Open Data Portals

Christine Meschede
Tobias Siebenlist
christine.meschede@hhu.de
tobias.siebenlist@hhu.de
Heinrich Heine University Düsseldorf
Düsseldorf, Germany

ABSTRACT

Since the adoption of the United Nations' Sustainable Development Goals (SDGs) in 2015 – an ambitious agenda to end poverty, combat environmental threats and ensure prosperity for everyone – some effort has been made regarding the adequate measuring of the progress on its targets. As the crucial point is the availability of sufficient, comparable information, open data can play a key role. The coverage of open data, i.e., data that is machine-readable, freely available and reusable for everyone, is assessed by several measurement tools. We propose the use of open governmental data to make the achievement of SDGs easy and transparent to measure. For this purpose, a mapping of the open data categories to the SDGs is presented. Further, we argue that the SDGs need to be tackled in particular at the city level. For analyzing the current applicability of open data for measuring progress on the SDGs, we provide a small-scale case study on German open data portals and the embedded data categories and datasets. The results suggest that further standardization is needed in order to be able to use open data for comparing cities and their progress towards the SDGs.

CCS CONCEPTS

• **Applied computing → E-government**.

KEYWORDS

Open Data, Sustainable Development Goals, Smart Cities, Open Governmental Data, Standards, DCAT-AP

ACM Reference Format:
Christine Meschede and Tobias Siebenlist. 2019. Open Urban Data and the Sustainable Development Goals: A small-scale Case Study on German Open Data Portals. In *dg.o 2019: 20th Annual International Conference on Digital Government Research (dg.o 2019), June 18–20, 2019, Dubai, United Arab Emirates.* ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/3325112.3325269

dg.o 2019, June 18–20, 2019, Dubai, United Arab Emirates

ACM ISBN 978-1-4503-7204-6/19/06…$15.00
https://doi.org/10.1145/3325112.3325269

1 INTRODUCTION

In 2015, The United Nations adopted the Sustainable Development Goals (SDGs) - 17 goals "to end poverty, protect the planet and ensure prosperity for all" [27, para. 1]. Sustainable development is defined as a "development that meets the needs of the present without compromising the ability of future generations to meet their own needs" [30, p. 37]. These needs originate from different perspectives, including social, environmental and economic aspects. To reconcile sustainability and development is an ambitious task. While, on the one hand, there is the need to sustain nature, environment, and culture, on the other hand, society and economy seek for development [15]. Sustainable development can hence be seen as a compromise between these interest groups [8].

The SDGs are an attempt to unite the various foci into one agenda (Table 1). In contrast to former similar strategic plans like the Millennium Development Goals, the agenda does not only focus on evolution in developing countries but demands action of all countries, even though different objectives shape different societies regarding social, economic and environmental goals [23]. The 17 SDGs include 169 more concrete targets. The interpretation of the targets can differ depending on the corresponding country. For example, SDG 1 includes the aim to reduce poverty according to national definitions.

An ongoing challenge is the detailed assessment of the SDGs and their targets all over the world. It has been agreed

on 232 indicators to measure the progress of each goal. The Inter-agency and Expert Group on SDG Indicators (IAEG-SDGs) developed the global indicator framework and classified each indicator into three tiers depending on methodological performance and availability of data [29]. The crucial aspect for defining the state-of-the-art concerning each indicator is the availability of sufficient information. Moreover, "the indicator framework should not only define what to measure but also how to measure it" [7, p. 566]. Data that is machine-readable, as well as available and reusable for everyone, i.e. open data, can hence play a key role in assessing the SDGs [6].

Table 1: Sustainable Development Goals [27]

SDG	Title
1	No poverty
2	Zero hunger
3	Good health and well-being
4	Quality education
5	Gender equality
6	Clean water and sanitation
7	Affordable and clean energy
8	Decent work and economic growth
9	Industry, innovation and infrastructure
10	Reduced inequalities
11	Sustainable cities and communities
12	Responsible consumption and production
13	Climate action
14	Life below water
15	Life on land
16	Peace, justice and strong institutions
17	Partnerships for the goals

Several initiatives and organizations focus on reporting on the progress of the SDGs on a global scale. For example, the SDG National Reporting initiative aims at facilitating information-sharing on the SDGs among policymakers, data managers, and international organizations [24]. The World Bank provides global development data that is free and openly accessible. One of their projects is the Atlas of Sustainable Development Goals, a detailed report and visualization of the SDGs' targets [31]. Other initiatives evaluate the general availability of data and statistics for sustainable development, often compared across different countries. The Open Data Inventory (ODIN) by the non-profit, non-governmental organization Open Data Watch monitors the coverage and openness of official statistics regarding social, economic and environmental information and ranks several countries in this regard [18].

Open data, and in particular open governmental data is being published in an increasing number of countries and their subordinate administrative units. Reference is made to the Open Definition, which states that the data must be available and usable freely and without restriction [20]. Furthermore, organizations stand behind the idea of open data and this as a basis for free knowledge. The Sunlight Foundation has published ten principles of open data [26], and the Open Knowledge Foundation has devoted itself to this topic in its Open Data Handbook [19]. In the recent years there has been a change in administration with regard to openness towards citizens. This movement is called open government [14], and its beginning is often associated with a memorandum from former US President Barack Obama in 2009 [16] in which he calls on administrations to act openly and transparently and thus also make datasets available. Also, more and more countries are creating legal requirements for this type of data publication, usually within the framework of freedom of information laws [25].

Various methods have been developed to measure the dissemination of open data in a country and thus create a basis for comparisons between countries. Such indices are presented in the following section. A relationship with the SDGs can thus be established at national level. Some categories of open data can be directly assigned to targets, while others are slightly more difficult to assign. However, open administrative data is not only published at the national level but also at subordinate levels. This data is thus also available at the city or municipal level, which allows a more detailed examination of the development and condition of these levels.

In this contribution, we propose the use of open governmental data and related measurement methods to make the achievement of SDGs easy and transparent to measure. For this purpose, a mapping of the categories or indicators to the targets is aimed at, which reflects the relationships between SDGs and open data. Since open data can also be found at lower administrative levels, this opens up a new possibility: the consideration of SDGs at the urban and municipal level. National evaluations reflect only the overall condition, without going into individual particularly good or bad examples. Citizens are increasingly aware of the environment in which they live and want to participate in the development of their city [4]. However, it is not always easy to participate as a citizen. Be it that the administration is not yet so open or that ideas and starting points within the framework of a pursuable plan are missing. Information and Communication Technology can act as a facilitator of such plans and foster citizen participation, as "the use of modern computer technology can contribute significantly to the empowerment of citizens through the collection, processing, enrichment and presenting information implicitly contained in open datasets" [11, p. 125]. Others believe that open data can also contribute

to foster aid effectiveness [10]. The SDGs can serve as a template for citizen participation. These are general goals that improve living together in the city for each and set out a plan. We want to take this fact into account by not only using open datasets to measure the achievement of the objectives but also by looking at them, in particular at the regional level. The resulting comparisons of cities and municipalities can serve as an incentive to improve the situation and as an impetus, e.g. competitions to improve one's city.

2 OPEN URBAN DATA, CITIZEN PARTICIPATION AND SUSTAINABLE DEVELOPMENT

Today, more than half of the world's population lives in cities. The United Nations estimated that by 2050 the number of urban residents will grow even more up to two thirds [28]. Hence, cities play a significant role in achieving sustainable development. This was also acknowledged by the United Nations by embedding a goal on sustainable cities and communities (SDG 11). Referring to this goal as the *urban sustainable development goal (USDG)*, Klopp and Petretta [9] argue that an ongoing challenge for the SDGs is the poor availability of urban data. In contrast, opening datasets at the city level can contribute to innovations and value-added city services [17],[12],[13] and thereby may facilitate sustainable development. Through the transformation to a smart city, factors such as participation, equity, mobility, and fairness can be improved [1].

The Open Data Charter Measurement Guide [2, p. 4] directly addresses this issue: "It is important to note that this guide is equally useful to assess government open data initiatives on a regional or at a city level." In the developing process of this guide and the corresponding mapping, also subordinate levels were thought of. While the above quotation is only about open data initiatives, it can be transferred to the SDGs on a regional level. Besides the actual embedding in the urban sustainable development goal – SDG 11 – cities are linked to every other SDG as well, as the addressed obstacles cannot be tackled without cities.

Data and their evaluations are required to assess the achievement of the SDGs and to calculate the scores of the corresponding indicators. These are often collected through surveys or expert interviews and are then reviewed. This can also be done at the city or the municipal level. Ideally, the data can be found in the open data portals or can be requested via them. They are therefore freely available. Validation from the administration's point of view usually already takes place when or even before publishing data on a portal. Other experts can be citizens who are familiar with their city and can review the data records published by the administration. This creates a new form of citizen participation that deals with the processing of open data to achieve higher goals.

Hence, the opening of datasets has enormous potential to foster citizen awareness. Especially at the city level, several initiatives can be recognized that contribute to enhanced participation in the matter of a community. For example, citizens collaboratively build particulate matter sensors on their own for measuring the air quality in their city themselves[1]. The data is then embedded in maps for visualizing the air quality. Besides, all data records are stored in an open format and can be retrieved at any time. Through Citizen Science projects like this, initiated by the Open Knowledge Foundation Germany[2], citizens are sensitized regarding environmental and social issues that are part of the SDGs and become producers of open data themselves. Issues that have to be considered in such approaches are data quality and privacy as well as the long-term impact of the produced data [3]. Nonetheless, what can be achieved with such citizen-driven projects can be seen in the OpenStreetMap[3]. This is a map that is continuously updated and expanded by volunteers from all over the world and is available to everyone freely and openly. Similar activities on other topics may produce records created by citizens for their fellow citizens that provide information about their city that would otherwise not exist. In this context SDG 17 also plays an important role. The partnerships for achieving the goals can also be related to national and municipal data providers as well as to citizens who collect data themselves, help with their evaluation, processing, and visualization and can thus directly participate in shaping processes. Citizen participation can, therefore, be seen as part of one of the objectives and thus also as a supportive means of achieving the other goals.

In the following section 3, we introduce measurement tools for the coverage and openness of open data and describe our mapping approach of the underlying data categories with the SDGs. The result of this mapping is presented in the subsequent section 4. Then, we concentrate on the results of our small-scale study regarding open datasets in German cities (section 5). Finally, we discuss the overall findings in section 6.

3 MATERIALS AND METHODS

In 2015, the Open Data Charter was founded by open data experts from governments, organizations, society and the private sector. It has been adopted by 54 national and local governments around the world and comprises the following six principles:

(1) Open By Default
(2) Timely and Comprehensive
(3) Accessible and Usable

[1] https://luftdaten.info/en/home-en/
[2] https://okfn.de/en/
[3] https://www.openstreetmap.org/

(4) Comparable and Interoperable
(5) For Improved Governance & Citizen Engagement
(6) For Inclusive Development and Innovation

According to the Open Data Charter [2] there exist five main measurement tools evaluating the prevalence of open government data: the Open Data Barometer (ODB), the Global Open Data Index (GODI), the OECD OURdata Index (OURdata), the Open Data Inventory (ODIN) and the European Open Data Maturity Assessment (EODMA). Three of these – ODB, GODI and ODIN – provide a classification into several data categories, which were chosen for further analysis. For better comprehensibility, the three open data measurement tools are presented in the following.

Open Data Barometer (ODB). The ODB is a collaborative project by the World Wide Web Foundation [32]. It evaluates the readiness for open data initiatives, the implementation of open data programs and the impact that open data is having on business, politics and civil society. Covered data categories range from national statistics to international trade data and crime statistics. Beside secondary data, the ODB uses an expert survey for evaluation. In 2016, Great Britain ranked first according to the ODB. Results and datasets are available under a Creative Commons license at opendatabarometer.org

Global Open Data Index (GODI). GODI was developed by Open Knowledge International and measures the openness of Open Government Data from a civic perspective, which means data should be useful for the public [21]. In the 2016 version, GODI aims explicitly at achieving comparability between datasets. Experts review data concerning several categories and indicators with different weightings are used. With an overall score of 90%, Taiwan scored best in 2016. The results are available in different formats under a public domain data license at index.okfn.org/download.

Open Data Inventory (ODIN). ODIN was initiated by the non-profit, non-governmental organization Open Data Watch and assesses 180 countries regarding the coverage and openness of official national statistics [18]. In contrast to the previously presented tools, ODIN explicitly analyses data categories that are oriented towards progress on the SDGs. By default, all categories and used indicators are weighted equally but can be customized by users. In 2017, the highest overall score was achieved by Denmark, whereas in 2016, Sweden occupied the first rank. All results and datasets are available under a Creative Commons license at odin.opendatawatch.com.

The Open Data Measurement Guide [2] provides the methodology for the mapping of open data indicators against the six principles of the Open Data Charter. This framework is a basis for our mapping, in which we use the categories from the three measurement tools and link them to the SDGs. In a first step, all data categories of GODI, ODB, and ODIN were sighted manually and checked for overlaps and similarities (Figure 1). For each tool, the latest available version was used for the analysis. In a second step, the concrete targets of the SDGs were analyzed regarding the applicability of the aforementioned data categories.

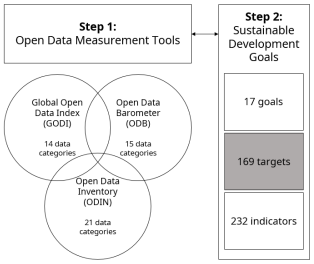

Figure 1: Mapping Open Data Categories with SDGs

In addition to the mapping, our goal is to compare the discussed open data categories with the actual availability of open datasets in cities. As a starting point, we concentrated on German large cities (cities with at least 100,000 inhabitants) and their open data portals. For the introduction of open data, the size of the city or municipality does not play an important role. The approach we present is therefore not limited to large cities but can be applied equally to municipalities of any size. The big cities in this analysis only serve as a starting point for our investigation. Germany is in the upper midfield in country comparisons such as those of the Open Data Barometer or the Open Knowledge Foundation Index. More and more municipalities provide open data portals or are joining forces to store data records in portals. Some pioneers started implementing open data and open government measures at a very early stage and share their experiences with other municipalities. The country comparisons mentioned show that the situation has improved compared to previous years and that the existence of open data is developing further. However, much remains to be done to catch up with the leading open data nations. Regarding German large cities, there are 79 cities that meet this criteria. For these, we checked the availability of an open data portal on the city level with the help of the Open Data Atlas[4], which lists the current status of open data portals in German-speaking countries. We identified 27 out of the initial 79 cities having an open data portal. In a next step, we retrieved each portal and investigated the embedded data categories. Therefore, we listed each category and the number of available datasets.

[4]http://opendata.tursics.de/

Four cities were excluded from the analysis because the listed portals either did not represent a general open data portal (Münster and Braunschweig concentrate on geodata), or the concrete number of datasets could not be identified (Berlin and Mannheim), ending up with a final set of 23 cities. The presented results refer to available data as of January 29^{th}, 2019.

4 OPEN DATA MEASUREMENT TOOLS AND THE SUSTAINABLE DEVELOPMENT GOALS

Table 2 presents our findings on the overlaps of three open data measurement tools concerning their covered data categories. For each category, the SDGs are noted whose targets address aspects of this category.

Some of the categories are directly linked to an SDG. For example, the category health is embedded with a corresponding goal (SDG 3: good health and well-being). Both ODB and ODIN evaluate countries regarding the provision of health sector datasets. ODIN further differentiates between health facilities, health outcomes, and reproductive health. Also, vital statistics like birth and death rates are of importance for measuring progress on SDG 3. Similarly to health, gender equality received its own SDG 5. In contrast, actual gender statistics are only considered by ODIN. Representative datasets provide information on violence against women, the proportion of women in government or data on child marriages.

SDG 4 addresses quality education for all. Among its targets (4.1) is the intention to ensure primary and secondary education for all girls and boys. Beside SDG 4, education plays a significant role for economic growth and is hence also embedded in SDG 8, target 8.1, which aims at reducing the proportion of youth not in employment, education or training. The Open Data Barometer evaluates countries regarding their performance on education data. Similarly, a category on education outcomes is integrated into the Open Data Inventory. On the contrary, the Global Open Data Index does not include a category regarding education statistics.

The Open Data Inventory is the only measurement tool integrating a particular category on poverty and income statistics, thus directly addressing SDG 1. Income statistics are further deemed relevant for SDG 11, which addresses the reduction of inequalities within and among countries.

Environmental data is handled differently in the three measurement tools. ODB includes a general category on environmental statistics which entails data like carbon emissions, particulate matter, and deforestation. With air and water quality, GODI provides two more specific categories. ODIN on the other hand again combines water and air quality to one category on pollution. Besides, ODIN further evaluates data coverage and openness regarding land, resource, and energy use. Environmental data does not only play a role for goals directly linked to the environment but rather is vital for several targets, which becomes apparent in the long list of SDGs listed in Table 2 for this data category. Apparently, water quality is directly linked to safe and affordable drinking water and hence to SDG 6. Similarly, the connection to SDG 14 (life below water) and SDG 15 (life on land) is self-evident. But as water and air quality are essential for health, a target of SDG 3 addresses the reduction of deaths and illnesses from air, water, and soil pollution. What's more, industry and innovation are also linked to environmental factors. SDG 9, therefore, includes, inter alia, increased use of clean and eco-friendly technologies. The same holds true for cities, which shall reduce their environmental impact (SDG 11). ODIN further covers a category on the built environment which includes the proportion of people with access to water and sanitation and is hence strongly linked to SDG 6. Besides, this category can also be mapped to SDG 11, as housing quality indicators pertain to safe and sustainable cities and human settlements.

Some open data categories cannot be dedicated directly to one single SDG, but are anchored on a lower level in different targets, e.g., legislation (GODI and ODB) or government spending (GODI, ODB, and ODIN). Legislation data plays a significant role for measuring progress on sustainable development, as "governments at all levels have had to develop an understanding of sustainable development and to consider how it should be integrated into everyday decisions and actions in their respective regions, countries, and municipalities" [22, p. 1104]. For the SDGs, legislation data was deemed relevant in particular for the goals 5, 8, 10, 14 and 15. For example regarding gender equality (5), one of the targets (5.c) aims at adopting "policies and enforceable legislation for the promotion of gender equality and the empowerment of all women and girls at all levels." Similarly, SDG 14 entails a target on implementing the national law for the conservation and sustainable use of oceans and their resources. Further appearances of legislation information among the targets can be found in SDG 8 regarding labor rights, in SDG 10 regarding the elimination of discriminatory laws and SDG 15 regarding national legislation for the prevention or control of invasive alien species.

Similar to legislation data, information on government spendings is anchored in several targets. For example, among the targets of SDG 2 is the increasing investment in rural infrastructure and agricultural research to enhance agricultural productive capacity in developing countries. All three measurement tools provide a corresponding category. While ODIN covers a general category on government finance which includes actual revenues and expenditures, GODI and ODB distinguish between government budget and spending; GODI further incorporates a separate category for procurement.

Table 2: Open data categories and related Sustainable Development Goals

GODI	ODB	ODIN	SDGs
Social statistics			
		Population & vital statistics	3
	Education	Education facilities Education outcomes	4, 8
	Health	Health facilities Health outcomes Reproductive health	3
		Gender	5
	Crime statistics	Crime & justice	16
		Poverty & income	1, 10
Economic statistics			
National statistics	National statistics	National accounts	8
		Labor	8, 9
		Price indexes	
Budget Spending Procurement	Budget Government spend	Government finance	1, 2, 11, 16, 17
Company register	Company registration		
		Money & banking	
	International trade	International trade	2, 10
		Balance of payments	
Environmental statistics			
Land ownership National maps Administrative boundaries Locations	Land ownership Mapping data	Land use	9
		Resource use	12
		Energy use	7
Water quality Air quality	Environment	Pollution	3, 6, 9, 11, 14, 15
		Built environment	6, 11
	Public Transport		9, 11
Other statistics			
National law Draft legislation	Legislation		5, 8, 10, 14, 15
	Contracting		
Election results	Election results		

5 OPEN DATA PORTALS IN LARGE GERMAN CITIES

In order to apply open data for measuring progress on the SDGs in cities, the prevalence of appropriate data is essential. With our small-scale case study we analyzed the situation in German large cities as a starting point. In Germany, only about one third (27 out of 79) of the large cities provides an open data portal according to the Open Data Atlas. Most cities show similarities in the selection of their data categories but differ immensely in the number of datasets provided for these categories.

One reason for this is the lack of a mandatory standard for the data categories. Work has been carried out on guidelines to develop a common national metadata standard. At the time of this decision, however, the first open data portals already existed, which differed in various points such as data categories. Subsequently, the existing German portals

or also international portals were taken as examples for the introduction of new open data portals, if municipalities had decided to implement open data.

At the European level, a common metadata standard (DCAT-AP[5]) was created which, among other things, provides for a list of 14 categories for data records (Table 3). However, this is not a strict requirement, but can be adapted to national differences. In such an adaptation, mapping to the DCAT-AP specifications should also take place, so that cross-country compatibility and comparability is ensured. These local adaptations should contain these mappings directly, and they should also be made available on the European Data Portal. The specifications for the controlled vocabulary in DCAT-AP, which covers different areas, are available in all languages of the European Union, so that uniform national terms and spellings are also standardized.

Table 3: Data Categories proposed by the DCAT-AP metadata standard

Abbreviation	Title
AGRI	Agriculture, fisheries, forestry and food
ECON	Economy and finance
EDUC	Education, culture and sport
ENER	Energy
ENVI	Environment
GOVE	Government and public sector
HEAL	Health
INTR	International issues
JUST	Justice, legal system and public safety
OP_DATPRO	Provisional data
REGI	Regions and cities
SOCI	Population and society
TECH	Science and technology
TRAN	Transport

In June 2018, the German adaptation DCAT-AP.de[6] was established as a formal exchange standard for open general administrative data. Implementation in open data portals is already underway and is scheduled for 2019. The maintenance manual for DCAT-AP.de stipulates that data at the national, federal and municipal level will comply with this standard. This will lead to the fact that the mapping in Table 2 may be subject to additions and maybe changes, but will subsequently be applied to all open data portals in all countries that operate their portals according to the DCAT-AP standard. Further, a mapping of Table 2 with the SDGs to the proposed categories of DCAT-AP will be performed. For regional peculiarities, the mapping of the local adaptations (for Germany, for example, DCAT-AP.de) is available, so that a two-stage assignment exists here.

In particular, this point in the various mappings shows how important common standards are in these subject areas in order to achieve evaluations and the best possible comparability. Only when a metadata standard such as DCAT-AP is mandatory in open data portals on all hierarchical levels of a country can a comparative analysis be carried out without very time-consuming and error-prone intellectual processing.

This does not exclude national peculiarities, as the local adaptations of DCAT-AP show. Each country can define its standard based on this meta-standard, which is, however, compatible with DCAT-AP and has a mapping. For cross-country comparisons, these local characteristics and mappings are to be included in the evaluation when it comes to checking whether the individual objectives have been achieved.

Table 4 shows the data categories for each city with the number of open datasets for each topic. Some categories directly match with the open data categories presented in the mapping of the previous section. For example, education, health, environment and legislation can be found in the Open Data Barometer. Taking a look at the number of datasets for these categories, we can observe some differences between the analyzed cities. For example, with 2802 datasets, Hamburg provides by far the largest amount of data for the category environment. Thereby, most sets consist of data on environmental measurements (2600), followed by public plans (133) and geodata (67). In comparison, all other cities provide between one and 39 datasets for this topic, resulting in a median of 7. For the category education, we could find the highest amount of datasets in Moers (31). In contrast, four cities do not provide any open data on education. For this category, the median value is 4. With a median of 1 and 0, respectively, data in the categories health and legislation is rarer. Most datasets for these topics could be found on the open data portals of the cities Hamburg with 19 datasets for the category health and Düsseldorf with 17 datasets on legislation.

Another category that can be found both in our mapping (ODIN) and in the open data portals of the examined cities is population data. Considering the median of 14, data for this category is the most prevalent among the German cities. Only three cities do not provide data for this category. The largest amount of datasets (57) could be found on the portal of Cologne.

The numbers from the currently available evaluation show that such an evaluation can function and that sufficient data are available to implement measurement of the objectives. However, at this point we cannot say anything about the quality of the data and the possible values for the achievement

[5]https://joinup.ec.europa.eu/release/dcat-ap/12
[6]https://www.dcat-ap.de/

Table 4: Availability of categories and datasets in German open data portals

	HH	M	K	F	D	L	BO	W	BN	KA	GE	C	AC	HAL	KR	FR	HRO	MH	P	WAN	UL	J	MO
Population	13	42	57	12	16	0	16	2	61	12	20	0	29	14	6	45	13	1	4	0	44	15	52
Education	2	2	14	1	14	0	4	3	24	5	4	8	0	1	4	19	13	0	3	0	19	8	31
Transportation	80	33	50	6	11	0	7	22	63	2	2	5	3	4	1	12	22	3	3	1	17	9	46
Geography and Geology	45	12	72	8	17	0	5	52	66	3	6	0	6	22	21	29	42	8	10	8	20	0	26
Culture	12	11	28	4	15	11	5	10	33	9	18	0	5	0	12	28	35	0	1	2	38	4	35
Environment	2802	1	25	2	20	18	3	4	39	3	2	6	2	31	1	7	25	1	1	9	22	38	7
Politics	13	3	44	37	19	0	32	7	31	5	2	0	21	6	1	87	24	6	7	4	22	2	10
Health	19	1	5	2	4	0	1	1	4	2	0	0	1	0	1	6	12	0	0	0	11	0	5
Social	10	5	18	5	26	37	25	2	20	10	15	8	1	0	9	7	9	0	2	0	28	0	1
Economy	6	14	3	6	4	29	0	3	0	9	7	0	3	2	0	37	41	0	1	0	24	9	0
Legislation	3	0	3	0	17	0	0	8	3	1	0	0	0	0	0	3	1	0	0	3	7	0	0
Public Administration	12	1	26	0	10	37	8	11	75	10	2	2	11	4	1	14	44	0	0	6	19	8	81
Infrastructure, Construction and Housing	2997	4	16	12	29	12	5	25	78	9	40	0	11	11	4	17	66	0	12	2	24	38	24
Consumer Protection	0	0	0	0	0	0	0	1	0	2	0	0	0	0	0	0	16	0	0	0	7	0	0
Sum	6014	129	361	95	202	144	111	151	497	82	118	29	93	95	61	311	363	19	44	35	302	131	318

of the objectives, as further mappings and investigations are necessary. Also, the numbers in the categories for the different cities are very unevenly distributed. The focus in this analysis was on German cities and even of these only a small proportion has open data portals. This clearly shows that many municipalities still have to provide for the introduction of open data in their open government strategies so that we can use open governmental data as the basis for a comprehensive calculation of the achievement of the Sustainable Development Goals. The results also show a clear discrepancy among cities that already have open data portals. If we take SDG 4 (Quality Education) as an example and look at the data sets in the Education category, a comparison between Hamburg and Moers shows that there is a striking difference in the number of existing data sets (2 compared to 31). In future work on this topic, it will, therefore, be necessary to assign the SDGs at the data set level from their targets or indicators and thus establish a closer link to specific open data sets.

6 DISCUSSION AND CONCLUSION

The assignment of open data categories to SDGs proposed in this paper is the first step towards an alternative, transparent calculation of an index for the status of achievement of the goals. The basis here is open governmental data, which must be available for the calculation. This also implies that the absence of an open government culture in a country makes the assessment impossible.

As soon as one starts comparing values for different countries, the fundamental question of comparability arises. Are the values for the SDGs of the individual directly comparable or do other factors have to be taken into account? The country profiles currently generated by the UN have additional development metrics that can be used to better classify the results. Also, all countries are classified into different categories, so that comparability in these categories is possible if necessary. This would also have to be done in the case of an adjustment at the city level, so that, for example, the number of inhabitants or the GDP are included in such a calculation or comparison.

To get an estimation on the progress of the SDGs, a platform for merging global open datasets would be useful. In some countries, there are already national portals, which harvests the data from the municipal portals and thus offer an interface for searching for data records from all parts of the country. On the European level, there exists the European open data portal[7], which also harvests the national open data portal from Europe and thus allows a more global insight. This idea further thought could lead to the creation of a worldwide portal that enables the search for all data records in all open data portals. The difficulties lie in the use of different portal software, different interfaces for accessing the data and different handling of metadata. In addition to that, barriers in opening governmental data exist [5], that should not be disregarded. However, further standardization in these areas will make this goal achievable in the future. The mapping of the data categories to the SDGs is not solved, however. This requires additional categorization based on our proposed mapping. Ireland already launched a portal[8] which specifically lists datasets for the SDGs and the concrete indicators. Projects like this could further contribute to

[7] https://www.europeandataportal.eu/
[8] http://irelandsdg.geohive.ie/

adequately measuring progress on the SDGs. In order to ensure comparability with other countries and to apply this on the city level, it is necessary to develop common standards for categories and data attributes.

A mapping can be conducted in a separate application which generates reports based on the country profiles which are already created by the UN for the SDGs. However, this application accesses (automatically) the data source (the possibly available global open data portal), extracts and processes the data and uses it to generate characteristic values that can represent the achievement of the individual goals per country. As already mentioned, it is possible in this context not only to access national, aggregated data but also to carry out a measurement for individual regions, cities and municipalities. This means that interested and committed citizens can also find out how their city has achieved one or all of the 17 goals. The data on which the calculation is based can also be viewed, and thus the values can be traced. This leads to more transparency and thus to verifiability and acceptance of the calculated values and rankings.

In this paper, we proposed a mapping of the open data categories of three measurement tools to the UN Sustainable Development Goals. In this way, several overlaps and similarities of the open data categories and SDGs became apparent, though with varying intensity depending on the respective measurement tool. While ODIN explicitly covers categories for social, economic and environmental development, ODB and GODI do not have this focus. Therefore, in particular, the categories of ODIN could be linked to the SDGs. For some categories a direct equivalent is anchored in the SDGs, this holds true for example regarding health, education or gender. Some further open data categories can be applied to several development targets, e.g., legislation data or statistics on government spending. Still, the number of open data categories oriented towards sustainable development can be further extended, as there is an excellent potential of open data to measure the progress of each SDG and its targets.

Section 5 shows the use of a controlled vocabulary metadata standard for the data categories. The use of such standards is a next important step to achieve measurement of target achievement and comparability based on open data. A mapping between the mapping generated in this paper (Table 2) and the categories from DCAT-AP (Table 3) is needed. Furthermore, local standards and their mapping to DCAT-AP must be taken into account in cross-country analyses. The path via categories and category measurements, therefore, represents a multi-stage procedure that can be integrated into a process for data collection and data analysis by using standards. The next step is to examine the SDGs more closely and classify them into appropriate needs for datasets.

The Sustainable Development Goals can then be assigned more precisely to the categories from the various indices with the DCAT-AP categories. This is particularly true in cases where a target is assigned to more than one category, what then leads to a further step, which will take place after the introduction of DCAT-AP and the corresponding mapping. To evaluate the achievement of the SDG or make calculations for a degree of progress, a data basis is required. The data basis must be comparable for the survey of different municipalities, federal states or states. The next step is, therefore, the use or creation of model data catalogs, which create standardization. This involves not only the names and sorting of the datasets into categories but also which datasets should be available in each portal. That means, for example, that for each municipality there is a record of student numbers, which can be found quickly and easily. A further mapping will follow, namely that of the targets or indicators to the data records described in sample data catalogs. An iterative process is needed, as the sample data catalogs must also develop and evolve. Such a development is taking place in the federal state of North Rhine-Westphalia in Germany: with support and expertise from Vienna (Austria), such a model data catalog is being created and its use tested[9]. Mapping with Sustainable Development must keep pace and be adapted accordingly. In this way, the calculation of the individual goals will be possible successively with open datasets that can be found uniformly, are described and are available in a comparable form for different municipalities.

Further, for applying the presented mapping in order to measure progress on the SDGs in cities and, in particular, to compare cities with each other, the availability of datasets is essential. With our small-scale study we provide a first insight into the challenges that exist for this approach. Cities might provide datasets that are not comparable to each other. Currently, the prevalence of data categories in German cities adheres in some parts to the ODB, which in turn can be applied to the SDGs to some extent. The availability of actual datasets for the individual categories varies immensely. Further, in our study we could not evaluate the actual quality and suitability of existing datasets. This should be considered in further studies.

The SDG index is undergoing changes in its development. Sources for indicators change, new indicators are added, and some indicators are removed or replaced by others. There is, therefore, an additional challenge to take these dynamic changes into account and to adapt the framework to the changes in the different versions of the SDG index. To show changes over time, the different versions of the SDG index must be brought in line with the mapping and the data records for the respective year. The results obtained for the

[9]https://blog-smartcountry.de/wie-foerdern-wir-open-data-in-kommunen-der-naechste-schritt/

national level can be compared with the official country profiles of the SDG index to get an impression as to what extent equal results are achieved. In the event of deviations from the official results, it remains to be investigated on which data basis and according to which methodology both values were calculated. Differences in the values obtained make it possible to initiate a discussion about transparency in the calculation of the values. There are no comparison values for the subordinate levels. More empirical studies are to be carried out to what extent the calculated values correspond to the condition of cities.

This contribution represents a first step on the way to a framework that enables the use of open governmental data to measure the achievement of SDG targets. In doing so, the focus is placed in particular on the local reference. It does not remain the status quo that the objectives are only considered at national level. It is precisely through the availability of open datasets of cities and municipalities that these goals can be considered for an area that is manageable and where the effects can be experienced and shaped directly by the local citizens. At the same time, the 17 goals offer a template for possibilities for local citizen participation to make one's city more livable, fairer and better for one's fellow citizens and also for oneself.

A LIST OF ANALYZED OPEN DATA PORTALS

(HH)	Hamburg	http://transparenz.hamburg.de/open-data/
(M)	Munich	https://www.opengov-muenchen.de/
(K)	Cologne	https://offenedaten-koeln.de/
(F)	Frankfurt	http://offenedaten.frankfurt.de/home
(D)	Düsseldorf	https://opendata.duesseldorf.de/
(L)	Leipzig	https://opendata.leipzig.de/
(BO)	Bochum	https://www.bochum.de/opendata
(W)	Wuppertal	https://www.offenedaten-wuppertal.de/
(BN)	Bonn	https://opendata.bonn.de/
(KA)	Karlsruhe	https://transparenz.karlsruhe.de/
(Ge)	Gelsenkirchen	https://opendata.gelsenkirchen.de/
(C)	Chemnitz	http://portal-chemnitz.opendata.arcgis.com/
(AC)	Aachen	http://offenedaten.aachen.de/
(HAL)	Halle (Saale)	http://www.halle.de/de/Verwaltung/Online-Angebote/Offene-Verwaltungsdaten/
(KR)	Krefeld	https://www.offenesdatenportal.de/organization/krefeld
(FR)	Freiburg im Breisgau	https://fritz.freiburg.de/Informationsportal/configurator?scenario=OpenData
(HRO)	Rostock	https://www.opendata-hro.de/
(MH)	Mühlheim a. d. Ruhr	https://geo.muelheim-ruhr.de/open-data/13819
(P)	Potsdam	https://opendata.potsdam.de/pages/home/
(WAN)	Herne	https://www.herne.de/Stadt-und-Leben/Stadtfakten/Open-Data/
(UL)	Ulm	http://daten.ulm.de/
(J)	Jena	https://opendata.jena.de/
(MO)	Moers	https://www.offenesdatenportal.de/organization/moers

REFERENCES

[1] Simon Elias Bibri and John Krogstie. 2017. Smart sustainable cities of the future: An extensive interdisciplinary literature review. *Sustainable Cities and Society* 31 (2017), 183–212.

[2] Ana Brandusecu and Danny Lämmerhirt. 2018 [Jun 10, 2018]. Open Data Charter Measurement Guide Internet:. https://open-data-charter.gitbook.io/odcmeasurement-guide/

[3] Ling Jyh Chen, Yao Hua Ho, Hu Cheng Lee, Hsuan Cho Wu, Hao Min Liu, Hsin Hung Hsieh, Yu Te Huang, and Shih Chun Candice Lung. 2017. An Open Framework for Participatory PM2.5 Monitoring in Smart Cities. *IEEE Access* 5 (2017), 14441–14454.

[4] Anna Chiesura. 2004. The role of urban parks for the sustainable city. *Landscape and urban planning* 68, 1 (2004), 129–138.

[5] Peter Conradie and Sunil Choenni. 2014. On the barriers for local government releasing open data. *Government Information Quarterly* 31 (2014), S10–S17.

[6] Joel Gorin, Laura Manley, and Audrey Ariss. 2015 [Jun 04, 2018]. Sustainable Development Goals and Open Data Internet:. http://blogs.worldbank.org/ic4d/sustainable-development-goals-and-open-data

[7] Tomáš Hák, Svatava Janoušková, and Bedřich Moldan. 2016. Sustainable Development Goals: A need for relevant indicators. *Ecological Indicators* 60 (2016), 565–573.

[8] Robert W. Kates, Thomas M. Parris, and Anthony A. Leiserowitz. 2005. What is sustainable development? Goals, indicators, values, and practice. *Environment* 47, 3 (2005), 8–21.

[9] Jacqueline M. Klopp and Danielle L. Petretta. 2017. The urban sustainable development goal: Indicators, complexity and the politics of measuring cities. *Cities* 63 (2017), 92–97.

[10] Dennis Linders. 2013. Towards open development: Leveraging open data to improve the planning and coordination of international aid. *Government Information Quarterly* 30 (2013), 426–434.

[11] D A Lusa, R Dos Santos Rabello, and C R Cervi. 2015. Open smart city view - An architecture for open government data manipulation and presentation at city level. In *Proceedings of the 14th International Conference WWW/Internet 2015*. WWW/Internet, Maynooth, Greater Dublin, Ireland, 119–126.

[12] A. Mainka, S. Hartmann, C. Meschede, and W.G. Stock. 2015. Mobile Application Services Based Upon Open Urban Government Data. In *iConference 2015 Proceedings*. iConference, Newport Beach, California, USA, 1–15.

[13] A. Mainka, S. Hartmann, C. Meschede, and W.G. Stock. 2015. Open government: Transforming data into value-added city services. In *Citizen's Right to the Digital City: Urban Interfaces, Activism, and Placemaking*. Springer, Singapore, 199–214.

[14] Patrice McDermott. 2010. Building open government. *Government Information Quarterly* 27, 4 (2010), 401–413.

[15] National Research Council, Policy Division, Board on Sustainable Development. 1999. Our Common Journey. A transition toward Sustainability.

[16] Barack Obama. 2009. Memorandum for the heads of executive departments and agencies. *Presidential Studies Quarterly* 39, 3 (2009), 429.

[17] Adegboyega Ojo, Edward Curry, and Fatemeh Ahmadi Zeleti. 2015. A tale of open data innovations in five smart cities. In *48th Hawaii International Conference on System Sciences*. IEEE, Kauai, HI, USA, 2326–2335. https://doi.org/10.1109/HICSS.2015.280

[18] Open Data Watch. 2018 [Jun 02, 2018]. Open Data Inventory 2017 Methodology Report Internet:. http://odin.opendatawatch.com/Downloads/otherFiles/ODIN-2016-Methodology.pdf

[19] Open Knowledge International. 2017 [Jun 12, 2018]. Open Data Handbook Internet:. http://opendatahandbook.org/

[20] Open Knowledge International. 2017 [Jun 12, 2018]. The Open Definition Internet:. https://opendefinition.org/

[21] Open Knowledge International. 2017 [Jun 14, 2018]. Global Open Data Index. Methodology Internet:. https://index.okfn.org/methodology/

[22] Andrea Ross. 2010. It's time to get serious-why legislation is needed to make sustainable development a reality in the UK. *Sustainability* 2, 4 (2010), 1101–1127.

[23] Jeffrey D Sachs. 2012. From Millennium Development Goals to Sustainable Development Goals. *The Lancet* 379, 9832 (2012), 2206–2211. https://doi.org/10.1016/S0140-6736(12)60685-0

[24] SDG National Reporting Initiative. 2018 [Jun 10, 2018]. Facilitating information-sharing on SDG reporting Internet:. https://www.sdgreporting.org/

[25] Elizabeth Shepherd. 2015. Freedom of information, right to access information, open data: who is at the table? *The Round Table* 104, 6 (2015), 715–726.

[26] Sunlight Foundation. 2017 [Jun 14, 2018]. Ten principles for opening up government information Internet:. https://sunlightfoundation.com/policy/documents/ten-open-data-principles/

[27] United Nations. 2015 [Jun 04, 2018]. Sustainable Development Goals Internet:. http://www.un.org/sustainabledevelopment/sustainable-development-goals/

[28] United Nations. 2015 [Jun 04, 2018]. World Urbanization Prospects. The 2014 Revision Internet:. https://esa.un.org/unpd/wup/Publications/Files/WUP2014-Report.pdf

[29] United Nations Statistics Division. 2018 [Jun 10, 2018]. IAEG-SDGs. Inter-agency Expert Group on SDG Indicators Internet: . https://unstats.un.org/sdgs/iaeg-sdgs/

[30] WCED. 1987.

[31] World Bank Data Team. 2018 [Jun 13, 2018]. The 2018 Atlas of Sustainable Development Goals: an all-new visual guide to data and development Internet:. http://blogs.worldbank.org/opendata/2018-atlas-sustainable-development-goals-all-new-visual-guide-\penalty-\@Mdata-and-development

[32] World Wide Web Foundation. 2016 [Jun 10, 2018]. Open Data Barometer. ODB Methodology Internet:. https://opendatabarometer.org/2ndEdition/about/method.html