IDFI

Institute for Development of Freedom of Information

MAIN CHALLENGES AND GAPS OF GEORGIA'S PUBLIC DATA ECOSYSTEM

Tbilisi, 2022

Main Gaps and Challenges of Georgia's Public Data Ecosystem

Contents

Introduction

Public data plays an important role in improving public welfare, modernizing the public sector, and implementing governance reforms. In this regard, open data, which is freely available to any interested party and, most importantly, easy to process and use for various purposes, has particularly great potential. Open data makes it possible to identify and predict important trends and circumstances, as well as to ensure accountability of the government, improve services and create new innovative applications.

Along with the progress in terms of digital development and open government practices around the world, major emphasis has been made globally on improving open data ecosystems. Georgia is no exception to this trend, seeing the reforms in the latest years, despite many set-back and challenges faced in the process.

The Institute for Development of Freedom of Information (IDFI), by virtue of its mission and vision, has been closely engaged at various levels in most of the activities surrounding the improvement of accessibility to open data since its founding. Therefore the organization is in a good position to produce a comprehensive and contextual assessment of major legislative, institutional, and practical gaps/needs in terms of access to open data in Georgia.

The analysis aims to outline and cross-reference Georgia's results in various International surveys/reports in order to identify the main challenges and needs of the country in terms of data accessibility and management practices. Particular emphasis will be made on the [Global Data Barometer](#) (GDB), published in May 2022, for which IDFI was the [regional coordinator](#). For increased scope and credibility, the analysis will take into account the latest results of the [Open Data Inventory](#) (ODIN) and the Open [Data Maturity Report](#) (ODM). The GDB's approach to the assessment of data availability and accessibility can be considered the most comprehensive out of the three, meanwhile, ODIN is concerned with data provided by the National Statistics Offices (NSOs) and ODM mainly focuses on Government Open Data Portals.

Major Takeaways from the Global Data Barometer

The Barometer is a multi-dimensional and multi-layered study that assesses the state of data for the public good in 109 countries. An expert survey was conducted from May 2019 – May 2021 to create a new global benchmark that looks at data governance, capability, availability, use and impact of data. The barometer's findings can be explored and viewed through its core (e.g. governance and capability) or thematic modules (e.g. climate action, land, public procurement) as well as on a country level.

In the Eastern Europe and Central Asian region, for which the research phase was coordinated by IDFI, the highest scores were attributed to Ukraine. The diagram below shows the average score obtained for all GDB indicators and modules on a 100-point scale. In this regard, Georgia ranks fifth out of 12 countries in the region with 40.25 points, after Ukraine, Armenia, Kazakhstan, and Moldova. It is worth noting that the GDB does not prioritize ranking results, instead, it is more focused on contextual and in-depth analysis of existing data ecosystems.

Overall Total Scores
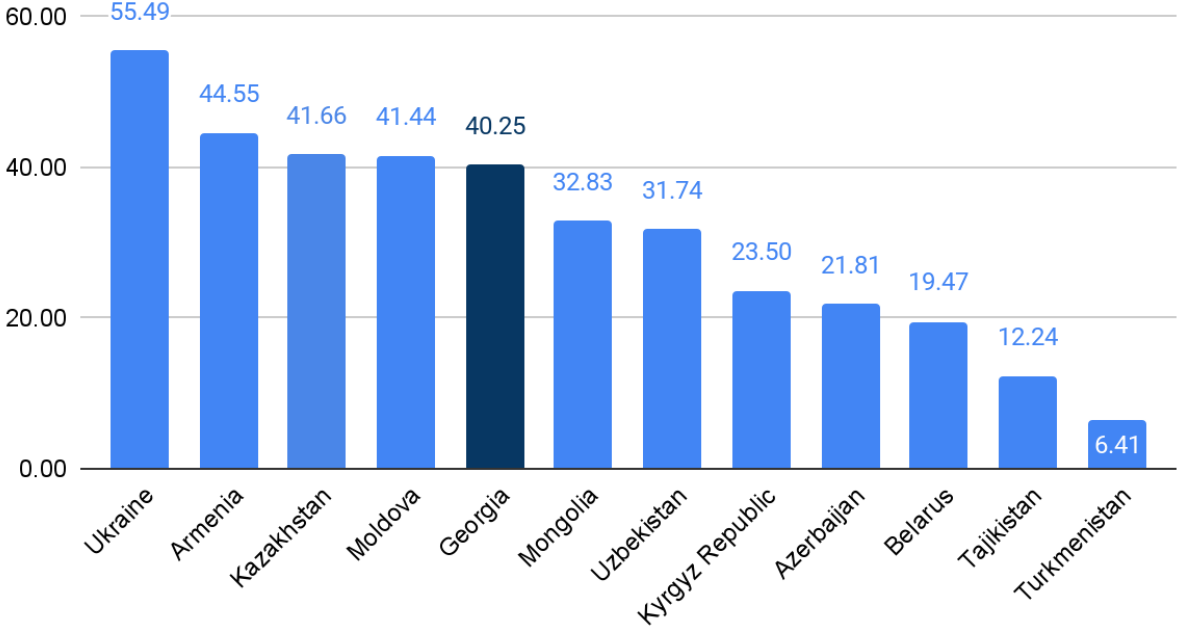
Below are presented the identified trends and challenges, as well as recommendations developed accordingly within the framework of the report:

- According to Data Governance indicators, Georgia ranks seventh in the region with 34.54 points. The results showed that it is important to establish/improve regulations, processes, and institutions to facilitate and protect data accessibility, and to strengthen existing efforts in this regard. Particular attention should be paid to the following requirements: a unified open data policy, a legal framework for data exchange, and a common standard for open data management.

- In terms of data capabilities, it is necessary to reactivate the work surrounding the existing government open data portal and regularly publish relevant datasets on it. In the initial stage, the involvement of all public institutions that collect and process data of high public interest is desirable in this process. Additionally, the need to develop data management capabilities at the local level was identified as one of the main challenges according to the survey.

- In the political accountability module, Georgia received one of the highest scores for the asset declarations indicator. The main need in this regard is open access to this data. The same module highlighted the need to improve public consultations as a mechanism for involvement in the

decision-making process and to establish mechanisms for data collection and processing in this area.

- There is a fee for obtaining complete data on land use in Georgia, which significantly hinders the availability of said data. In order to meet the open standards in this component, it is necessary to remove the Paywall.

- The state collects and publishes data on private companies, although it is also important to encourage the use of this data by the public. In addition, it is important to adopt and implement beneficial ownership transparency standards.

- In terms of public finance, it is essential that budget data be made available in bulk, in machine-readable formats. Additionally, the data needs to be disaggregated at the level of individual transactions, and the requirement to publish structured data should be reflected at the legislative level.

- It is important to collect real-time healthcare system capacity data in the country and publish it in an openly accessible format. Vaccination data should be made available in bulk, machine-readable format, and disaggregated by various indicators.

- In terms of climate data, it is important to start collecting and processing vulnerability data, both at the national and local levels.

GDB Indicators with the lowest scores in the Case of Georgia[1]	
Indicators	Score (out of 100)
Governance Module	
Data management frameworks	0
Data sharing frameworks	0
Open data policy	0
Capabilities Module	
Support for re-use	11
Sub-national capabilities	11

---

[1] See definitions and descriptions of each indicator in Annex I

Climate Action	
Vulnerability data	0
Health & Covid-19	
Real-time healthcare system capacity data	0
Land	
Gender and inclusion uses of data	4
Political Integrity	
Public consultation[2]	0
Lobbying data	16
Political integrity interoperability	0

Major Takeaways from the Open Data Maturity

The Open Data Maturity Report is an annual survey of European data portals. This report offers an extensive view into the developments in the open data environment in European countries, including the 27 EU Member States, EFTA countries (Iceland, Norway, and Switzerland) and Eastern Partnership countries (Georgia and Ukraine), as well as Montenegro and the United Kingdom. The report measures open data maturity in four main dimensions: policy, impact, portal, and quality. The latest edition was published in March 2022 and it overviews the overall data maturity scores of the 2021 assessment.

Georgia ranked last out of the 34 evaluated countries, in all four components of the ODM. The following issues were highlighted in the final report:

- Comparing overall open data maturity score in 2021 to previous editions, most countries show increased results while Georgia stagnates.

---

[2] Since May 2021, particular positive developments occurred regarding requirements of publication of consultation data, which might have positive affect on the countries performance in the future assessment.

- In Georgia, only a few public sector data providers contribute to the open government data portal. Furthermore, the databases published on the portal are not up to date.

- There is no evidence in Georgia of any mechanisms in place for measuring and assessing the extent of open data re-use. Additionally, There are no attempts at measuring the social, political and economic impact of data.

Based on the analysis of the ODM report and the results of Georgia, the following trends were identified and specific recommendations were developed[3]:

- At this stage, the Government of Georgia still has not developed an open data regulatory framework and national strategy. Because of this, the country has received low scores in the open data policy dimension. It is imperative to start working on these documents in a timely manner or to add open data related comittments to the updated version of the Open Government Partnership Action Plan, in order to clearly emphasize the importance of open data at the national level.

- It is important to develop a guidebook at the national level, to assist open data publishers in the disclosure process. A similar guidebook has already been introduced in other Eastern Partnership countries.

- In comparison to other countries, Georgia has made little effort to measure the impact of open data use, which has led to the lowest results in this dimension. Current trends in EU countries show that more and more attention is paid to monitoring and measuring the social and economic impact of open data. The introduction of this practice in Georgia will facilitate both the commercial use of open data and the promotion of civic initiatives.

- It is important to update the Georgian Open Data Portal, taking into account international good practices and important global trends, and incorporating improvements such as publishing data in more formats including API, diversity of published data, encouraging the use of databases and the creation of new services/applications by displaying them in the relevant section of the website. It is also important to systematically monitor the number of unique visitors and active users of the portal and make this information available publicly.

- One of the trends identified in the report is related to the increased importance of open data and information due to the COVID-19 pandemic. In response, many EU countries implemented various measures to improve access to data (awareness-raising campaigns, promotion of data use, development of data-driven platforms, etc.). In this regard, activity in Georgia is quite low and

---

[3] For more detailed analysis, please find IDFI's study on this issue:
https://idfi.ge/en/georgias_performance_in_the_open_data_maturity_2020_report

lags behind other Eastern Partnership countries. Therefore, it is necessary to ensure open access to any type of data primarily related to the Covid-19 pandemic.

Major Takeaways from the Open Data Inventory

The Open Data Inventory (ODIN) measures how complete a country's statistical offerings are and whether its data meet international standards of openness. Data assessed in ODIN must be official country data that are published on the national statistics office's (NSO) website or any other official country website that is linked to the NSO website.

Despite the fact that ODIN only evaluates the National Statistics Offices, in contrast with the GDB and the ODM index, some common denominators can still be identified. The latest edition covers data from 2020 and below is a table representing Georgia's rankings.

Rankings	Coverage	Openness	Overall
Global OUT OF 187	27th	32nd	31st
Western Asia OUT OF 18	2nd	5th	4th

Georgia ranked 2nd in the Western Asian Region and 27th in the World in terms of the coverage element, which evaluates the disaggregation level of data, availability of recent and historical data, and the availability of data at the subnational levels. It has to be noted that the latter component proved to be the most challenging, with most of the evaluated indicators lacking subnational disaggregation. Meanwhile, in the openness element, Georgia scored 5th in the region and 32nd in the World based on the assessments of the following components, machine readability, nonproprietary format, metadata availability, download options, and terms of use.

Data Category	Coverage	Openness	Overall
Population & vital statistics	70	90	80
Education facilities	30	90	60
Education outcomes	0	0	0
Health facilities	60	60	60
Health outcomes	30	90	60

Reproductive health	60	90	75
Food security & nutrition	63	80	72
Gender statistics	50	70	60
Crime & justice	50	90	70
Poverty & income	60	90	75
National accounts	75	100	89
Labor	60	90	75
Price indexes	75	100	89
Government finance	75	30	50
Money & banking	100	70	81
International trade	100	100	100
Balance of payments	100	70	81
Agriculture & Land Use	70	70	70
Resource use	75	70	72
Energy	100	70	81
Pollution	38	50	44
Built environment	50	70	60
All Categories	62	74	69

Types of data based on ODIN indicators that are less available in Georgia:

Data Category	Indicator	Description
Missing data categories		
Education Facilities	Education expenditures	The amount of money spent on education in a country in any currency. Disaggregated by school Stage, school type, and Expenditure Categories
Education Outcomes	Completion or graduation rate	Proportion of students from a cohort enrolled in a given grade at a

		given school year who study in the next grade in the following school year. Disaggregated by Sex, Age, School Stage, and School Type
	Competency exam results	The results for any standardized test administered in a country for math, science, or reading subjects or a combination of them. Disaggregated by Sex, Age, and School Stage
Health Outcomes	Immunization rate	The number of people who have been immunized against certain diseases per any number of inhabitants.
Reproductive Health	Maternal mortality rate	The number of maternal deaths during a given time period per live birth during the same time period.
Food Security & Nutrition	Prevalence of undernourishment	A percentage estimate of the proportion of the population whose habitual food consumption is insufficient to provide the dietary energy levels that are required to maintain a normal active and healthy life.
	Prevalence of moderate or severe food insecurity	The prevalence of moderate or severe food insecurity in the population, based on the Food Insecurity Experience Scale.
Required disaggregation by sex in the following data categories:		
Education outcomes	Enrollment rate	The number of students enrolled in a given level of education, regardless of age, expressed as a percentage of the official school-age population corresponding to the same level of education.
Health outcomes	Disease prevalence	The number of existing cases (new and old) from certain diseases at the end of a specific year.
Crime & justice	Homicide rate	The total count of unlawful death inflicted upon a person with the intent to cause death or serious injury per unit of population.

Labor	Employment distribution	The percentage distribution of those employed by their industry or occupation.

Overarching shortcomings of Georgia's national statistics practices:

Need for publishing more data in machine-readable formats:
10 ODIN indicators were available in datasets that were only made available in non-machine-readable format (PDF). In order for the data to be structured and readable by a computer, it should be accessible in (XML), (JSON), (CSV), (XLSX) or other formats. The datasets currently only available in PDF formats include Health expenditures, Data on land use or land cover, Data on protected lands, Data on timber harvests or deforestation, Energy consumption, Greenhouse gas emissions, and others.

Need for Publishing more comprehensive metadata:
40 ODIN indicators[4] were available in datasets that were not made available for all years with complete metadata. Many metadata standards can be used to meet ODIN criteria for complete metadata. Metadata must include the following three components: 1) Source agency of data, 2) Definition of indicator, and 3) Date that data was uploaded online.

Need for publishing more historical data as well as recent data:
11 ODIN indicators[5] were available in datasets that did not include data points for 3 of the last 5 years, and 10 ODIN[6] indicators were available in datasets that did not include data points for 6 of the last 10 years.

Need for publishing more subnational level data
38 ODIN indicators[7] were available in datasets that did not include data at the subnational geographic level. Disaggregation of indicators at the subnational level is one of the key elements to effectively monitoring the development of a country and ensuring certain groups of people are not left behind.

Conclusion

Despite the differing methodologies and varying scope of the analyzed reports, there were some gaps and challenges that presented themselves consistently. There is solid ground to argue that the pivotal shortcoming in terms of availability and accessibility of open data in Georgia is the absence of open data legislation. Although Georgia has undertaken several reforms related to open data in the scope of the

---

[4] See Annex II for a full list of indicators missing a metadata component
[5] See Annex III for a full list of indicators missing recent or historical data
[6] See Annex III for a full list of indicators missing recent or historical data
[7] See Annex IV for a full list of indicators missing subnational level data

Open Government Partnership (OGP) and subscribes to the IMF Standards for Data Dissemination, the country still lacks an enforceable and comprehensive legal act to regulate the disclosure of Government-held data.

ODIN and GDB both highlighted the data availability and accessibility shortcomings at the sub-national level. There is a need to develop data management capabilities locally, as well as the need to disaggregate centrally collected data in terms of municipal units. Another common issue emphasized by all three reports is the lack of effort and tools for measuring the social and economic impact of public data, which is crucial for the further development of data ecosystems and efficient prioritization of datasets.

Apart from the institutional and legislative gaps, certain specific datasets have been identified as difficult to access, undisclosed, or completely unavailable. The most problematic datasets include but are not limited to healthcare system capacity data, public consultation data, lobbying data, education expenditures, completion/graduation rates, immunization data, etc.

Lastly, The ODM results overviewed in this analysis highlight the urgent need to update [Georgia's open data portal](#). Consistent technical support, publication of new databases and regular updating of existing data are essential for the continued functioning of the portal. Furthermore, it is important to provide access to data through more open formats (especially API).

Three different international surveys of data accessibillity have identified significant challenges and needs for Georgia, although most of them are interrelated and can be addressed through the implementation of a comprehensive policy. The introduction of an open data policy and common standards based on international good practices will be an important step to overcome a significant part of the identified challenges. To this end, there is a need for a strong political will, which will ensure that decision-makers are aware of the importance, benefits, and key enablers for solid open data ecosystems in the country.

Annexes

Annex I

Indicators	Description
Governance Module	
Data management frameworks	To what extent do relevant laws, regulations, policies, and guidance provide a comprehensive framework for consistent data management and publication?
Data sharing frameworks	To what extent do relevant laws, regulations, policies, and guidance provide a comprehensive framework for data sharing?
Open data policy	To what extent do relevant laws, regulations, policies, and guidance provide a comprehensive framework for generating and publishing open data?
Capabilities Module	
Support for re-use	To what extent is there evidence that the government is providing support for data reuse?
Sub-national capabilities	To what extent do city, regional, and local governments have the capability to manage data effectively?
Climate Action	
Vulnerability data	To what extent is climate vulnerability information available as open data?
Health & Covid-19	
Real-time healthcare system capacity data	To what extent is information about the real-time capacity of the healthcare system available as open data?

Land	
Gender and inclusion uses of data	To what extent is there evidence that land data is being used to influence policy in the interests of equitable and inclusive land tenure and use?
Political Integrity	
Public consultation	To what extent is public consultation information available as open data?
Public consultation frameworks	To what extent do relevant laws, regulations, policies, and guidance provide a basis for collecting and publishing data generated through and about public consultation on rulemaking?
Lobbying data	To what extent do relevant laws, regulations, policies, and guidance provide a basis for collecting and publishing data on lobbying activities?
Political integrity interoperability	To what extent is political integrity data interoperable across different political integrity datasets, as well as other datasets associated with relevant information flows?

Annex II

Missing Metadata Component	Indicator
Definition	- Birth rate - Death rate - Number of schools or classrooms - Number of teaching staff - Enrollment rate - Number of health facilities - Diseases prevalence - Infant mortality rate or neonatal mortality rate - Under-5 mortality rate - Fertility rate - Contraceptive prevalence rate - Adolescent birth rate - Proportion of women in government, management or senior positions - Homicide rate - Crime rate

	- Poverty rate - Employment distribution - Data on land use or land cover - Commondity production - Energy consumption - Energy supply - Greenhouse gas emissions - Emissions of air or water pollutants - Number of rooms or bedrooms
Date	- Contraceptive prevalence rate - Prevalence of obesity - Prevalence of stunting - Prevalence of wasting - Proportion of women who are victims of physical, sexual, or psychological violence - Proportion of women who are victims of physical, sexual, or psychological violence - Data on child marriages
Date, Definition	- Health expenditures - Proportion of households with access to water - Proportion of households with access to sanitation - Number of rooms or bedrooms - Access to electricity - Data on housing construction materials
Source, Date, Definition	- Actual revenues - Actual expenditures - Data on fishery harvests

Annex III

Datasets that lack recent data	Datasets that lack historical data
- Health expenditures - Contraceptive prevalence rate - Prevalence of obesity - Prevalence of stunting - Prevalence of wasting - Proportion of women who are victims of physical, sexual, or psychological violence - Data on child marriages - Data on prison population - Emissions of air or water pollutants - Number of rooms or bedrooms - Data on housing construction materials	- Prevalence of obesity - Prevalence of stunting - Prevalence of wasting - Proportion of women who are victims of physical, sexual, or psychological violence - Data on child marriages - Data on prison population - Data on fishery harvests - Proportion of households with access to sanitation - Number of rooms or bedrooms - Data on housing construction materials

Annex IV

ODIN Indicator	Disaggregation Found
Population data	Sex, Marital Status, 5-Year Age Groups
Death rate	Sex
Number of schools or classrooms	School Stage, School Type
Number of teaching staff	School Stage, School Type
Enrollment rate	Sex, Age, School Stage, School Type
Number of health facilities	Facility Type
Number of beds or data on health care staff	Facility Type, Department Type, Staff Type
Health expenditures	Expenditure Categories
Diseases prevalence	Age, Sex, Disease Type
Infant mortality rate or neonatal mortality rate	Sex
Under-5 mortality rate	Sex
Fertility rate	None
Adolescent birth rate	None
Proportion of women in government, management or senior positions	None
Homicide rate	Sex Of Victim, Age Of Victim, Sex Of Perpetrator, Age Of Perpetrator, Victim/Perpetrator Relationship
Crime rate	Crime Type, Sex Of Victim, Age Of Victim, Sex Of Perpetrator, Age Of Perpetrator, Victim/Perpetrator Relationship
Data on prison population	Age, Sex, Sentencing Status
Poverty rate	None
Distribution of income by deciles or Gini coefficient	None

Gross domestic product (production approach)	Industry
Gross domestic product (expenditure approach)	Expenditure Categories
Gross domestic product (income approach)	Income Components
Employment rate	Sex, Age
Employment distribution	Industry, Occupation, Sex
Unemployment rate	Sex, Age
Consumer price index	None
Producers price index	None
Actual revenues	Revenue Source
Actual expenditures	Administrative Classification, Economic Classification, Functional Classification
Data on protected lands	None
Commondity production	Commodity Type
Data on fishery harvests	None
Data on mining or extractive activities	Activity Type
Water consumption	Sector
Greenhouse gas emissions	Gas Type
Emissions of air or water pollutants	Pollutant Type
Proportion of households with access to water	Water Supply Type
Proportion of households with access to sanitation	Sanitation Facility Type