

INTRODUCTION TO OPEN DATA

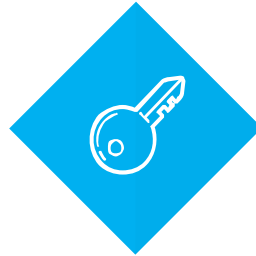
PART 1:

Definitions, benefits and challenges, and international standards and practices

Schedule

Topics and activities

Open Data 101
An overview of the Open Definition.



Benefits and Challenges
Balancing the economic, legal, and social benefits and challenges of opening data.



International Consensus
What has the global community done to support open data? Are there international standards?



Group Discussion
Group discussion followed by a question & answer session.



Open Data 101

What types of data should be open?



Mapping



Meteorological



Land Ownership



Budgets



Company Registration



Public Transit Timetables



Election Results



Public Contracts



National Statistics

Defining Open Data

Open data

[oh-puh n] [dat-uh]

noun

“Open data and content can be freely used, modified, and shared by anyone for any purpose.”

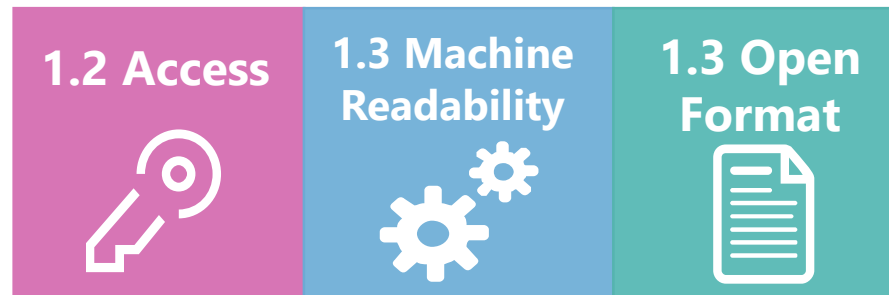
Source: The Open Definition

Two Dimensions of Data Openness

Legally Open



Technically Open



The Open Definition 2.1

There are four main criteria of openness:

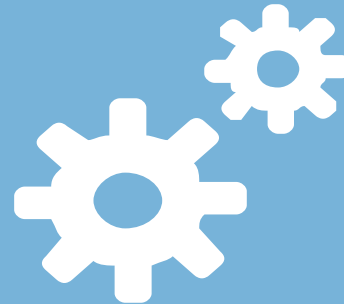
1.1 Open License



1.2 Access



1.3 Machine Readability



1.3 Open Format



1.1 Open License



“The **work** *must* be in the **public domain** or provided under an **open license**”

Examples of open licenses:

- Creative Commons Licenses (CC0, CC BY 4.0, CC BY 4.0 SA).
- Public Domain Dedication and License (PDDL)
- Canada’s Open Government License (Version 2)
- United Kingdom’s Open Government License (Version 3)

:

Creative Commons BY 4.0

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material
for any purpose, even commercially

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

1.2 Access



“The **work** *must* be **provided as a whole** and at no more than a reasonable one-time reproduction cost, and *should* be downloadable via the Internet **without charge**”

Main criteria:

- Data should be available in bulk download.
- Data should be free of charge.
- Data should be publicly available on the internet.

Common barriers to access

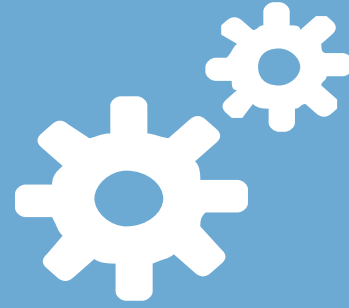
Required user login

Some websites require users to login to either access data or to download data. Both are restrictive measures and make the data no longer publicly accessible.

Lack of bulk download

Bulk downloads allow users to export data without having to download multiple files for data on the same or related indicators.

1.3 Machine Readability



“The **work** *must* be provided in a form readily processable by a computer and where the individual elements of the work can be easily accessed and modified”

Main Criteria:

- Data should be published in XLS, XLSX, CSV, XML or similar formats.

Different formats for different needs

Give users options

Not all users want data in the same format. CSV files are great for importing data into other systems, but XLSX files are easier for people to read.

Don't abandon PDFs

Data should always have a machine-readable format option for export, but this doesn't mean you should not publish in PDFs. Data in PDF reports often provides useful context to non-technical readers. Provide both if possible.

1.4 Open Format



“The **work** *must* be provided in an open format. An open format is one which places no restrictions, monetary or otherwise, upon its use and can be fully processed with at least one free/libre/open-source software tool.”

Main Criteria:

- Data should be published in non-proprietary format such as CSV, XLSX, PDF, DOCX, or similar.

:

What makes a format open?

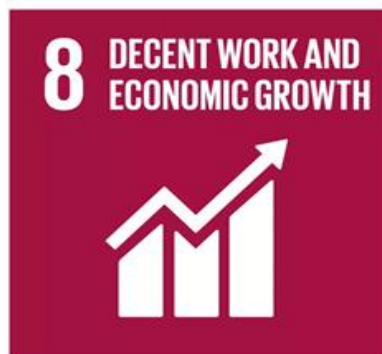
No price barriers

Data in a non-proprietary format requires users to purchase software to use it.

Fully readable by open source software

Data in open formats can be correctly read by a range of different software programs. Some closed formats are not readable by open source software, or only partially readable.

International Consensus



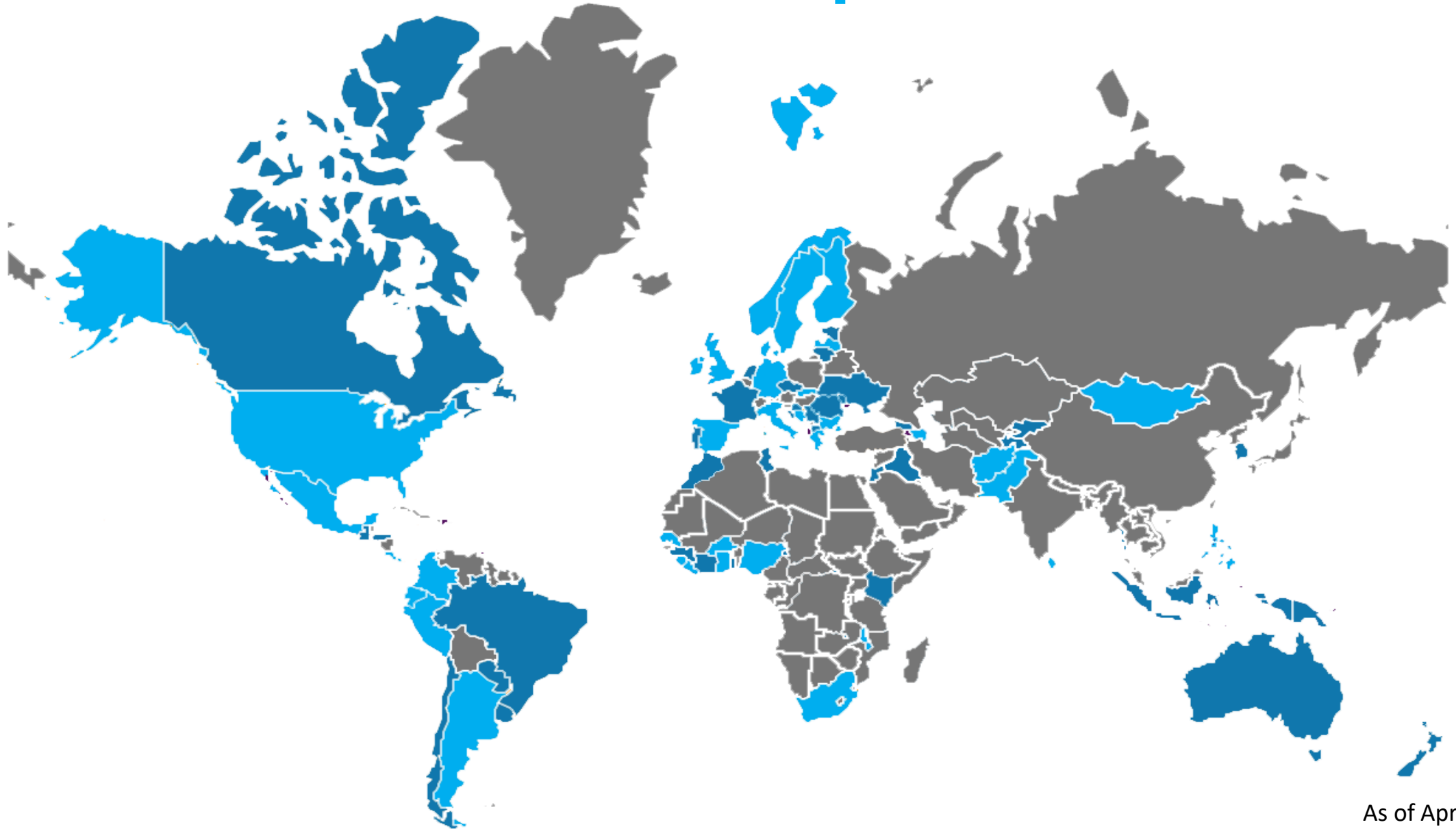


Open Data Seminar at
the 48th UNSC in March
2017

International Seminar on
Open Data for the SDGs in
South Korea in Sept 2017

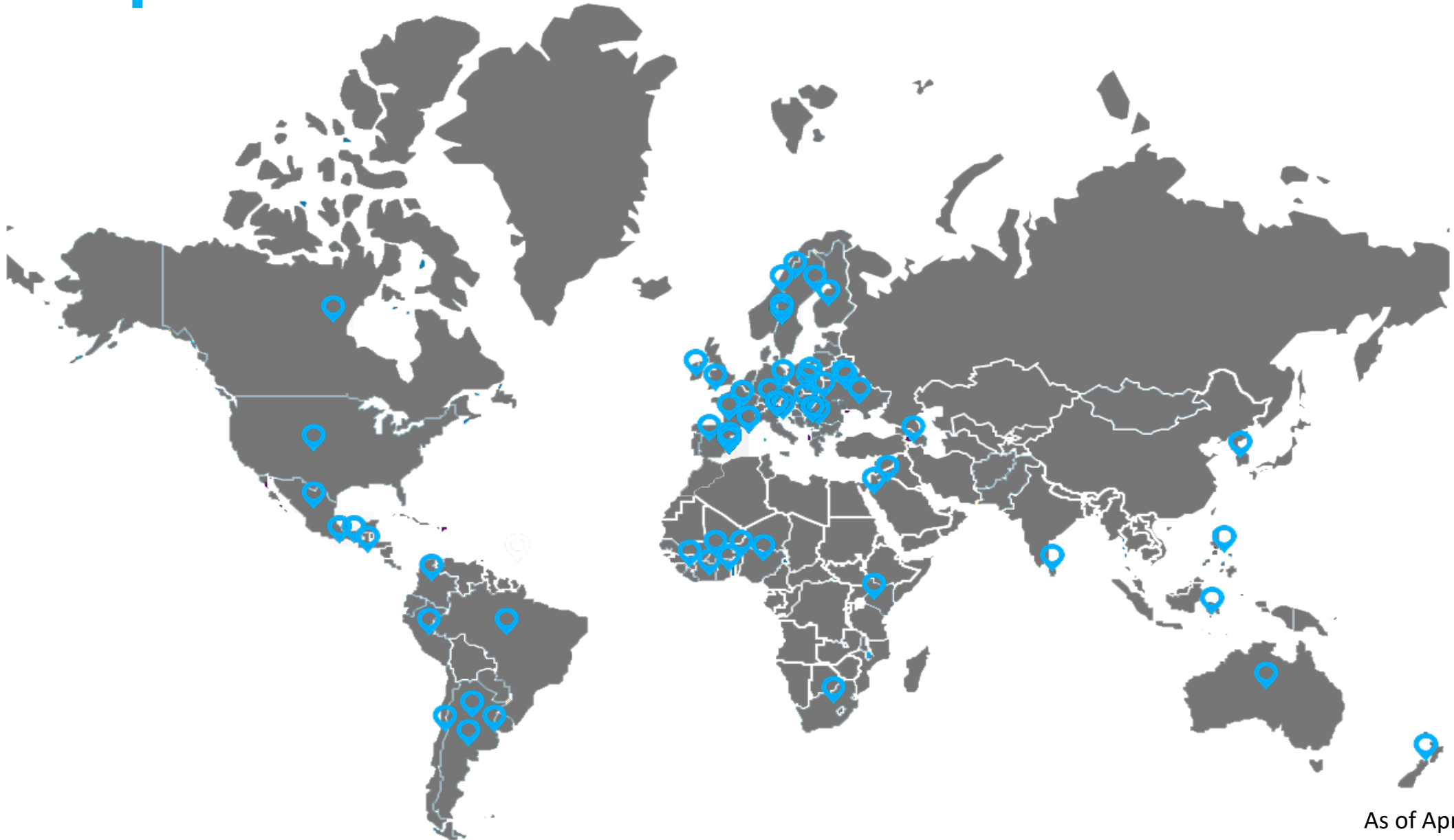
UNSD Open Data paper
presented at 49th UNSC
in March 2018

Open Government Partnership (OGP)



As of April 2019

OGP Open Data Commitments



As of April 2019

Open Data Charter



1. Open by Default



2. Timely and Comprehensive



3. Accessible and Usable



4. Comparable and Interoperable



5. For Improved Governance and Citizen Engagement



6. For Inclusive Development and Innovation

www.opendatacharter.net

Open Data Charter Adoption

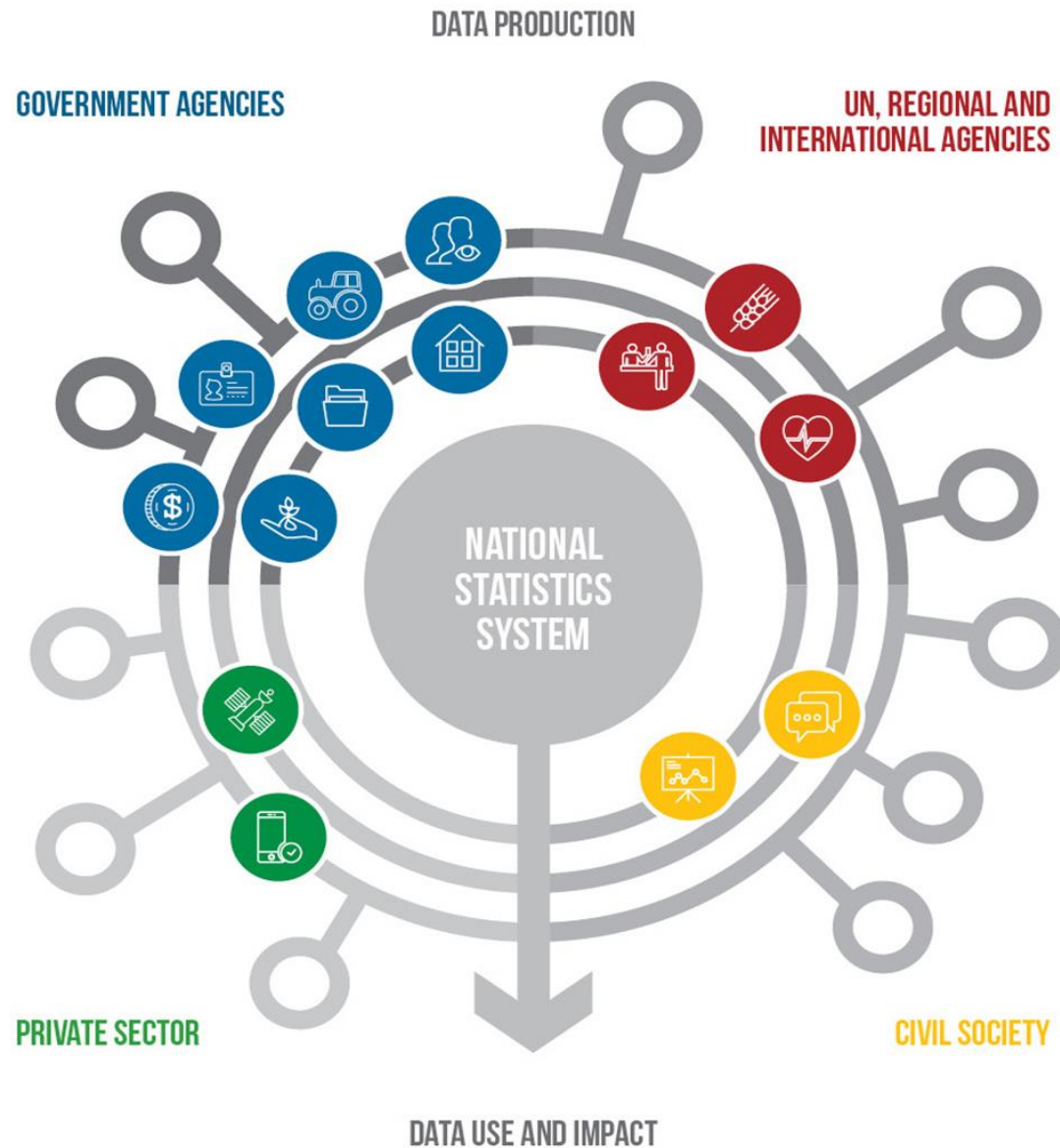


As of September 2018

There is a consensus.

National statistical systems can and should take the lead in an open data revolution.

Data Ecosystem



Uniquely positioned in the center of our expanding data ecosystem, national statistical systems will be intermediaries for change.

Benefits & Challenges

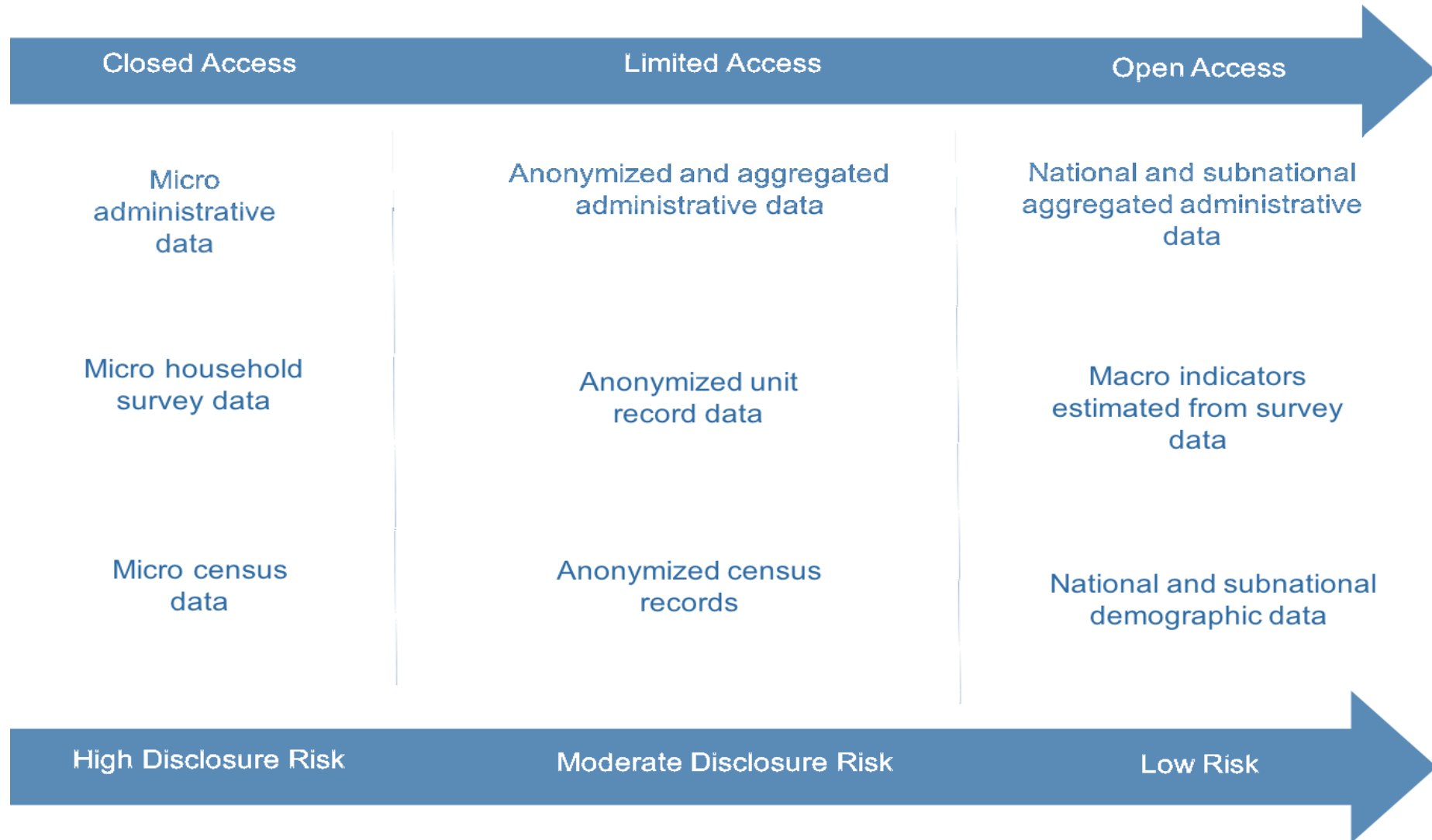
Balancing act



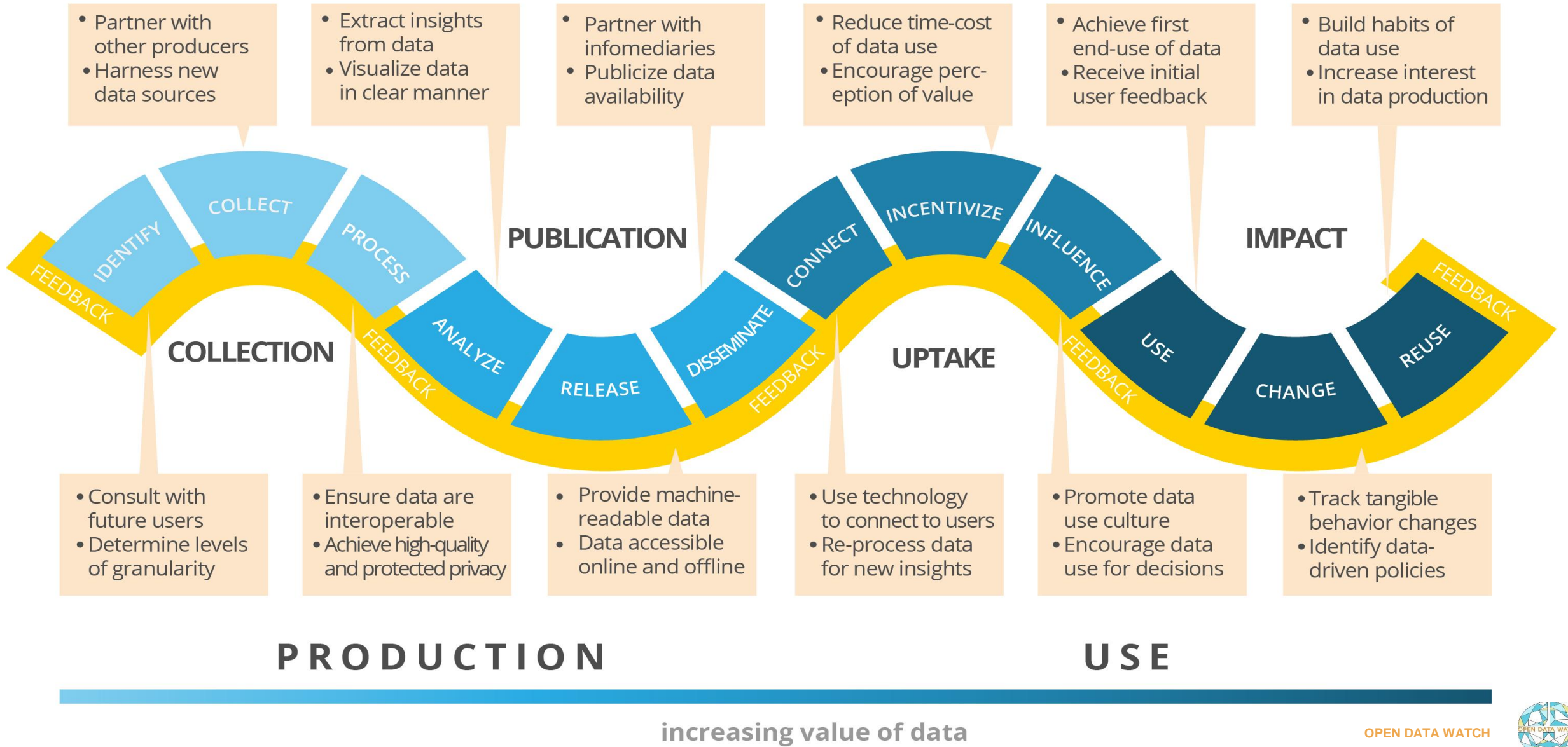
Benefits of open data

	Benefits	Challenges
Economic	<ul style="list-style-type: none">▪ Reduces costs associated with data requests▪ Increase accountability and decrease redundancy of services.▪ Create jobs and encourage entrepreneurship	<ul style="list-style-type: none">• Finding new revenue streams to replace revenue generated from selling data• Building the technical capacity of staff• Lack of technological resources• Extensive time investment to convert paper/PDF files to open formats
Other	<ul style="list-style-type: none">• Increase use and applicability of data.• Increase capacity to monitor development progress• Increase citizens access to information• Raises profile of National Statistical Office and other statistics producers	<ul style="list-style-type: none">• Lack of political will/support• Data sharing and interoperability• Legal barriers• Privacy and security

Data privacy spectrum



DATA VALUE CHAIN





MEASURING OPEN DATA

PART 2:

Open data indexes, methodology, and country results

Schedule

Topics and activities

Group Exercise: Analyzing open data gaps

Compare results on different open data measures for your countries. Draw conclusions and set priorities.



Measuring open data

What measures exist to track progress on open data? How do they differ?



Group Exercise: Discussion of results

Share highlights from small group discussions.

Measuring Open Data

OPEN DATA INVENTORY



**GLOBAL
OPEN DATA INDEX**



WORLD BANK GROUP

Open Data Readiness Assessment (ODRA)



OpenData
Barometer

Open Data Readiness Assessment (ODRA)

- A freely available tool developed by the World Bank's Open Government Data Working Group to assess the readiness of a government or agency to evaluate, design and implement an Open Data initiative.
- Assessments can be done by the agency, or agencies may request that the World Bank perform the assessment.
- Eight dimensions assessed, including policy/legal framework, institutional structures, demand for open data, and funding.

Bangladesh has not completed an ODRA

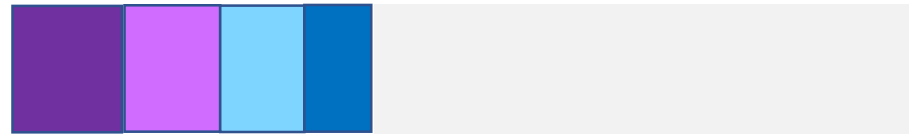
Open Data Barometer (ODB)

- Assessment conducted by the World Wide Web Foundation. It has three components it measures:
 - Readiness** for open data initiatives.
 - Implementation** of open data programmes.
 - Impact** that open data is having on business, politics and civil society.
- Assessments combine data from expert questionnaires, a government self assessment, and secondary indicators.
- The Open Data Barometer measures performance against the Open Data Charter principles.

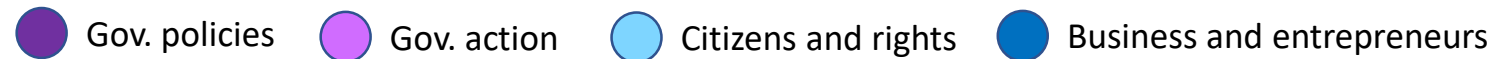
Open Data Barometer (ODB), 2017

Bangladesh's Overall Score: 11/100

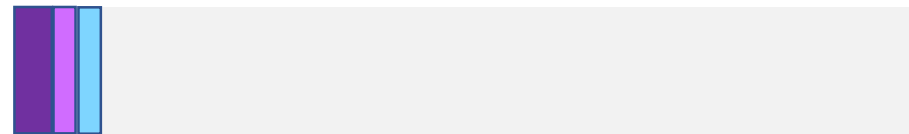
Readiness



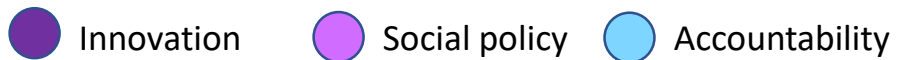
Score: 32/100



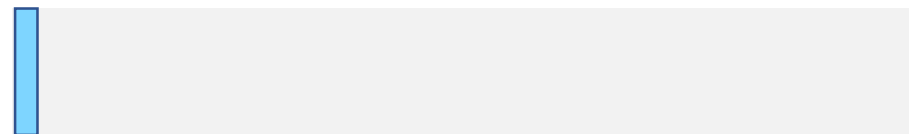
Implementation



Score: 6/100



Impact



Score: 2/100



Global Open Data Index (GODI), 2017/18

- Assessment conducted by Open Knowledge International that measures the openness of national government data. Use or impact is not assessed, only publication.
- The assessment looks for data across 14 categories, one which is national statistics. Datasets are measured against the open definition to determine degree of openness.
- Data is collected via crowdsourcing. Anyone can contribute data for the assessment.

Global Open Data Index (GODI), 2017/18

Bangladesh's Overall Score: 25/100

Openly Licensed



1 out of 15 datasets meet criteria

Machine Readable and Non-proprietary Format



1 out of 15 datasets meet criteria

Downloadable at Once (Bulk Download)



5 out of 15 datasets meet criteria

Up to Date



6 out of 15 datasets meet criteria

Publicly Available



5 out of 15 datasets meet criteria

Free of Charge



6 out of 15 datasets meet criteria

Global Open Data Index (GODI), 2017/18

Bangladesh's Overall Score: 25/100

Dataset	Score
Administrative Boundaries	100%
Government Budget	60%
National Laws	60%
National Statistics	45%
Procurement	45%
Land Ownership	45%
National Maps	15%
Draft Legislation	0%
Air Quality	0%
Weather Forecast	0%
Company Register	0%
Election Results	0%
Locations	0%
Water Quality	0%
Government Spending	0%

Open Data Inventory (ODIN) 2018/19

- Assessment conducted by Open Data Watch that measures the openness and coverage of government statistical data as published online.
- The assessment looks for data across 21 categories. Datasets are scored on how complete they are and how open they are (using criteria from the Open Definition).
- Data is collected by trained researchers and go through two rounds of internal review, plus a government review if they choose to participate.

ODIN Data Categories

Social

- Population & Vital Statistics
- Education Facilities
- Education Outcomes
- Health Facilities
- Health Outcomes
- Reproductive Health
- Gender
- Crime & Justice
- Poverty & Income

Economic

- National Accounts
- Labor
- Price Indexes
- Government Finance
- Money & Banking
- International Trade
- Balance of Payments

Environmental

- Land Use
- Resource Use
- Energy Use
- Pollution
- Built Environment

How ODIN Measures Coverage

- Data must be published publicly before we can assess openness.
- ODIN assessments looks for a predetermined set of indicators that are defined as fundamental to a country's statistical system. Those datasets are given scores based on the criteria below to measure their "completeness".

Coverage Criteria

(1) Are the data published and disaggregated?

(2) Are the data available over the last 5 years?

(3) Are the data available over the last 10 years?

(4) Are the data available at the first administrative level? (divisions)

(5) Are the data available at the second administrative level? (districts)

Measuring Coverage: Labor Statistics

Labor Statistics

Indicators

Employment rate

Employment distribution

Unemployment rate

Disaggregation

Sex, age

Sex, industry, occupation type

Sex, age

Coverage Questions

1. Are all indicators available and disaggregated?
2. Are the data available the last 5 and 10 years?
3. Are data available at the subnational level (divisions and districts)?

How ODIN Measures Openness

- Like other open data indexes, ODIN's scores openness against adherence to the open definition.

Coverage Criteria

(1) Are the data published in a machine readable format?

(2) Are the data published in a non-proprietary format?

(3) Can you download the data in bulk or custom selection?

(4) Are the data published with metadata?

(5) Is the terms of use policy for the data fully open?

Measuring Openness

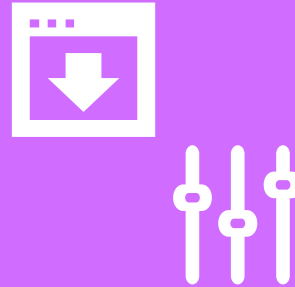
- Once data is found, ODIN scores how well the data adheres to the open data principles.

OPEN FORMATS



Are the data machine readable and in a nonproprietary format?

BULK/CUSTOM DOWNLOADS



Are the data downloadable in bulk or can a user customize the export file?

METADATA



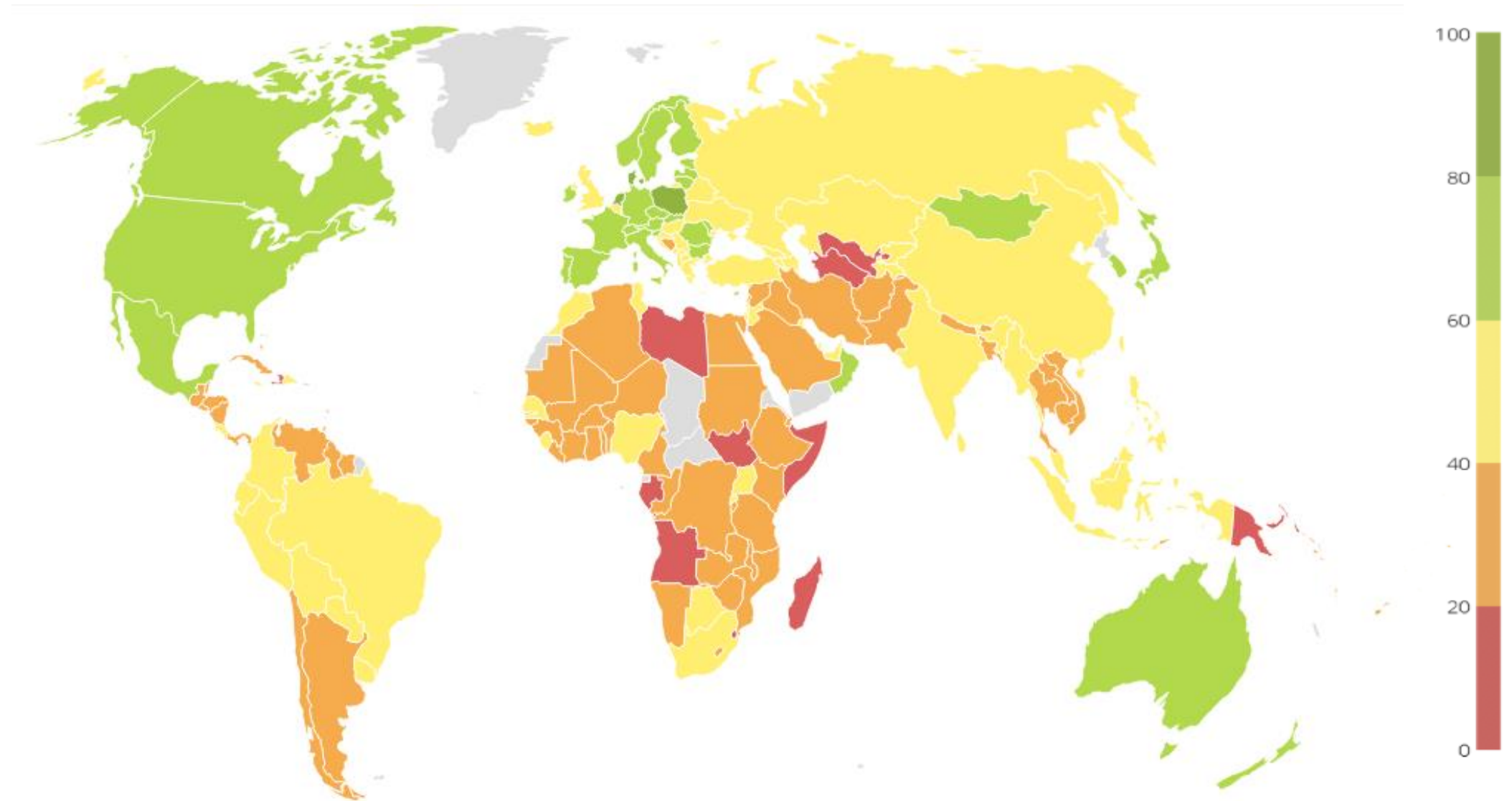
Are the data accompanied by metadata to help users understand the dataset?

TERMS OF USE



Are the data published under an open license? Does the license have restrictions?

ODIN 2018/19 Global Results



ODIN 2018/19 Results



Bangladesh

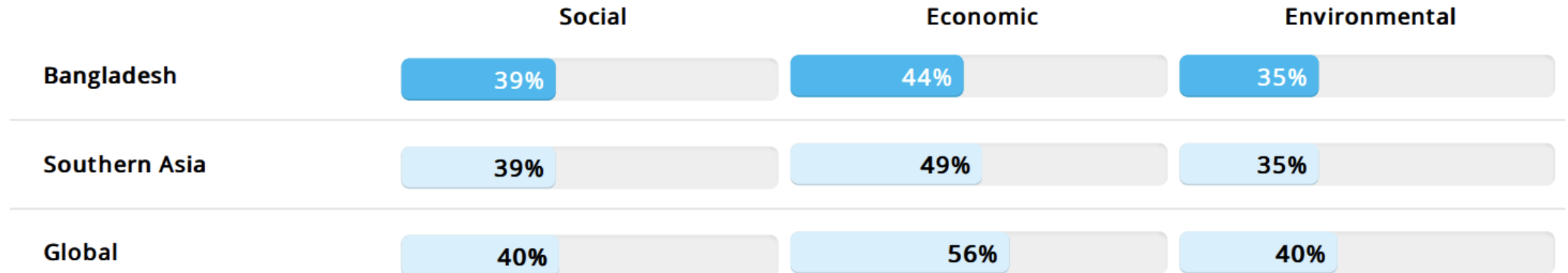
National Statistics Office of Bangladesh
<http://www.bbs.gov.bd>

105 RANK
out of 178

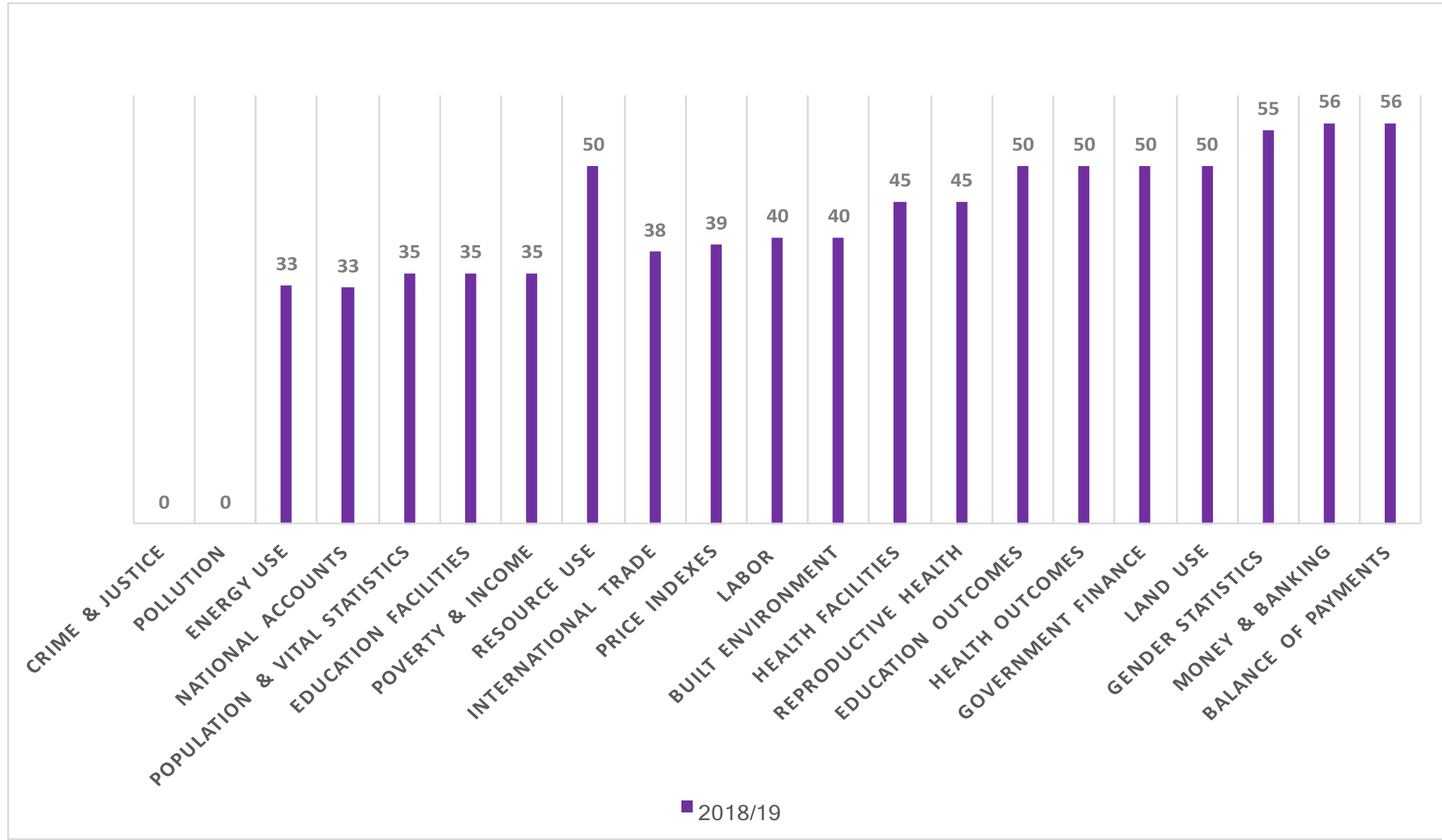
44 COVERAGE
subscore

35 OPENNESS
subscore

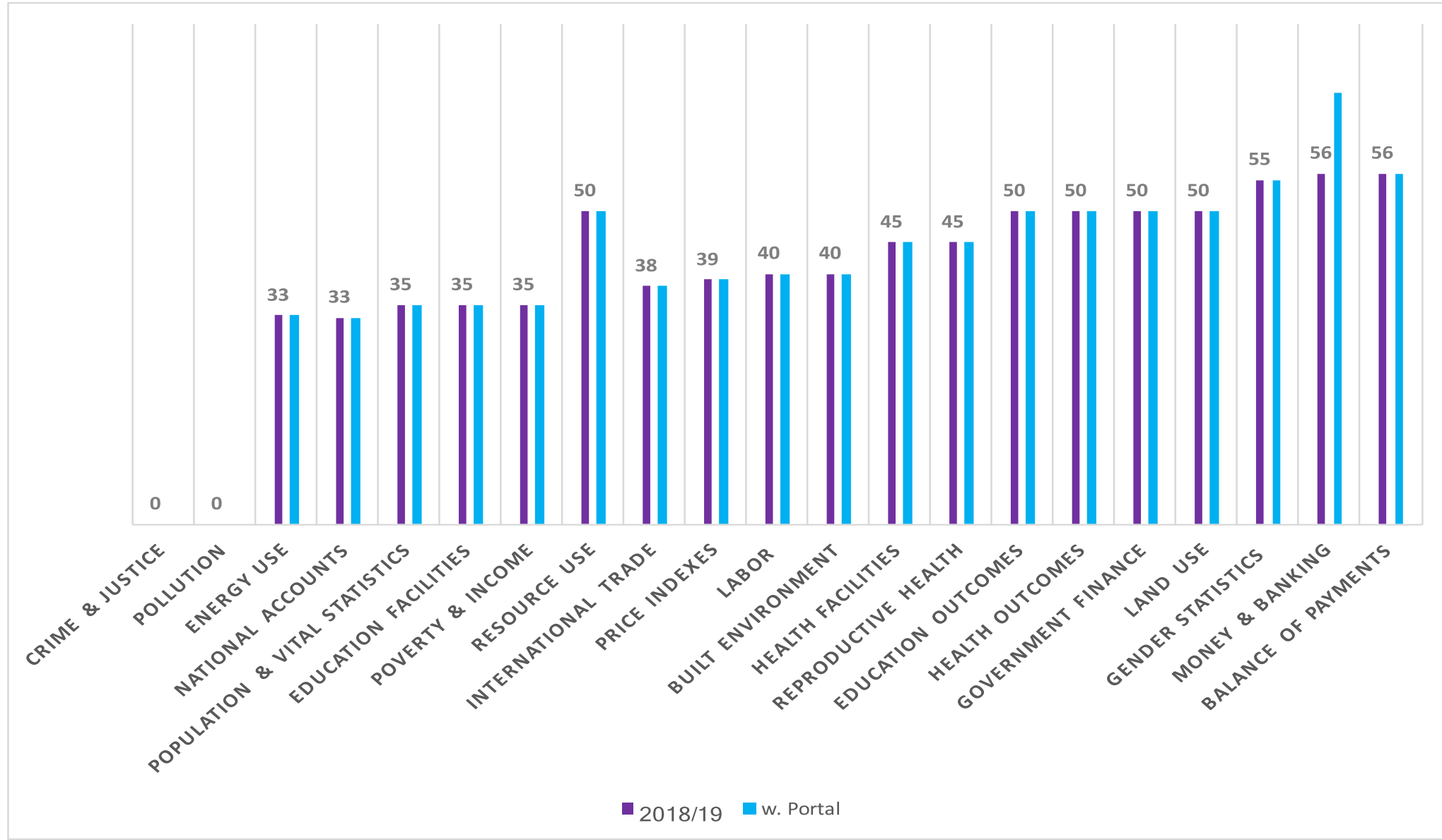
SCORES BY MAJOR DATA CATEGORY



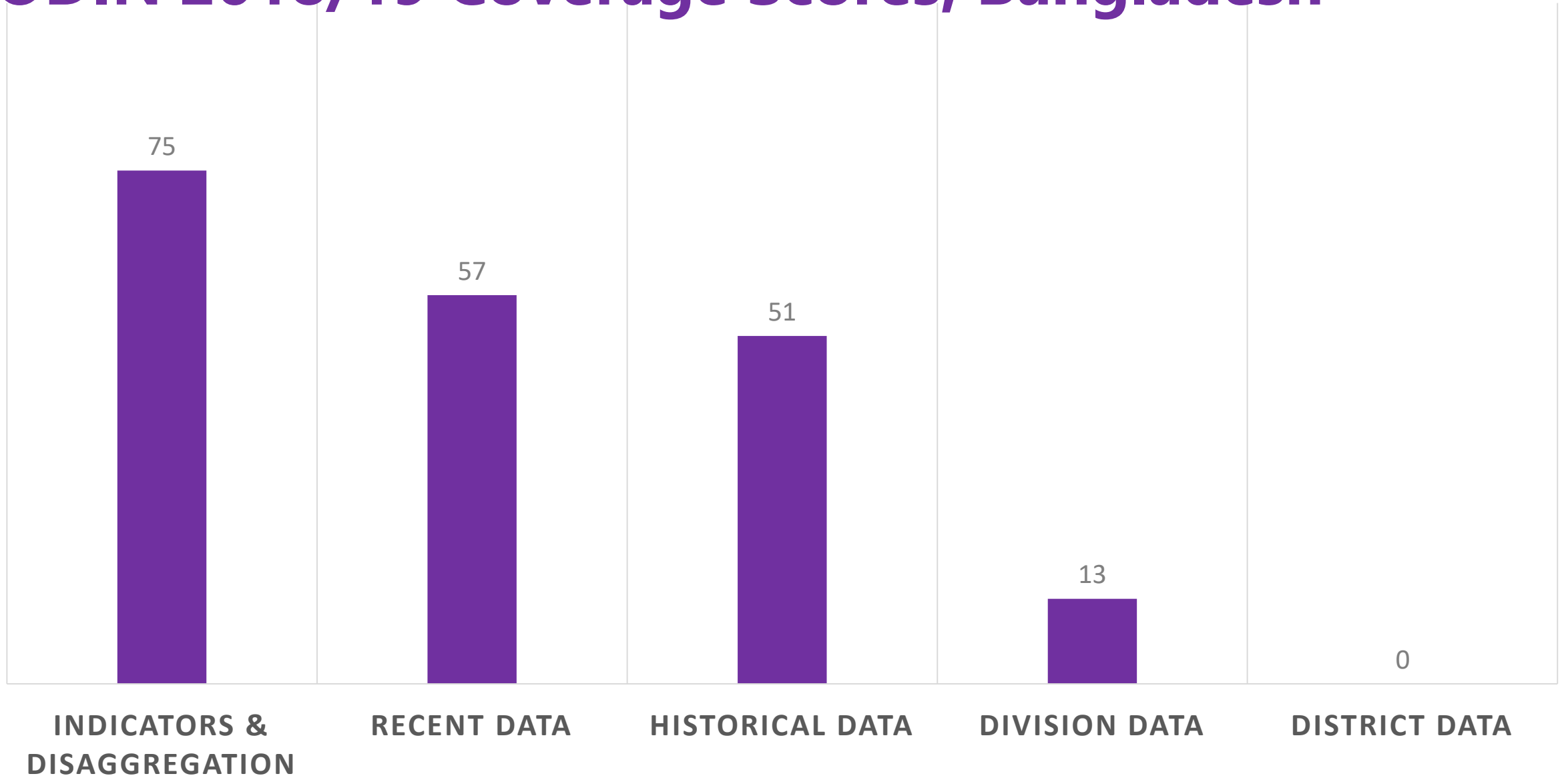
ODIN Scores by Category, Bangladesh



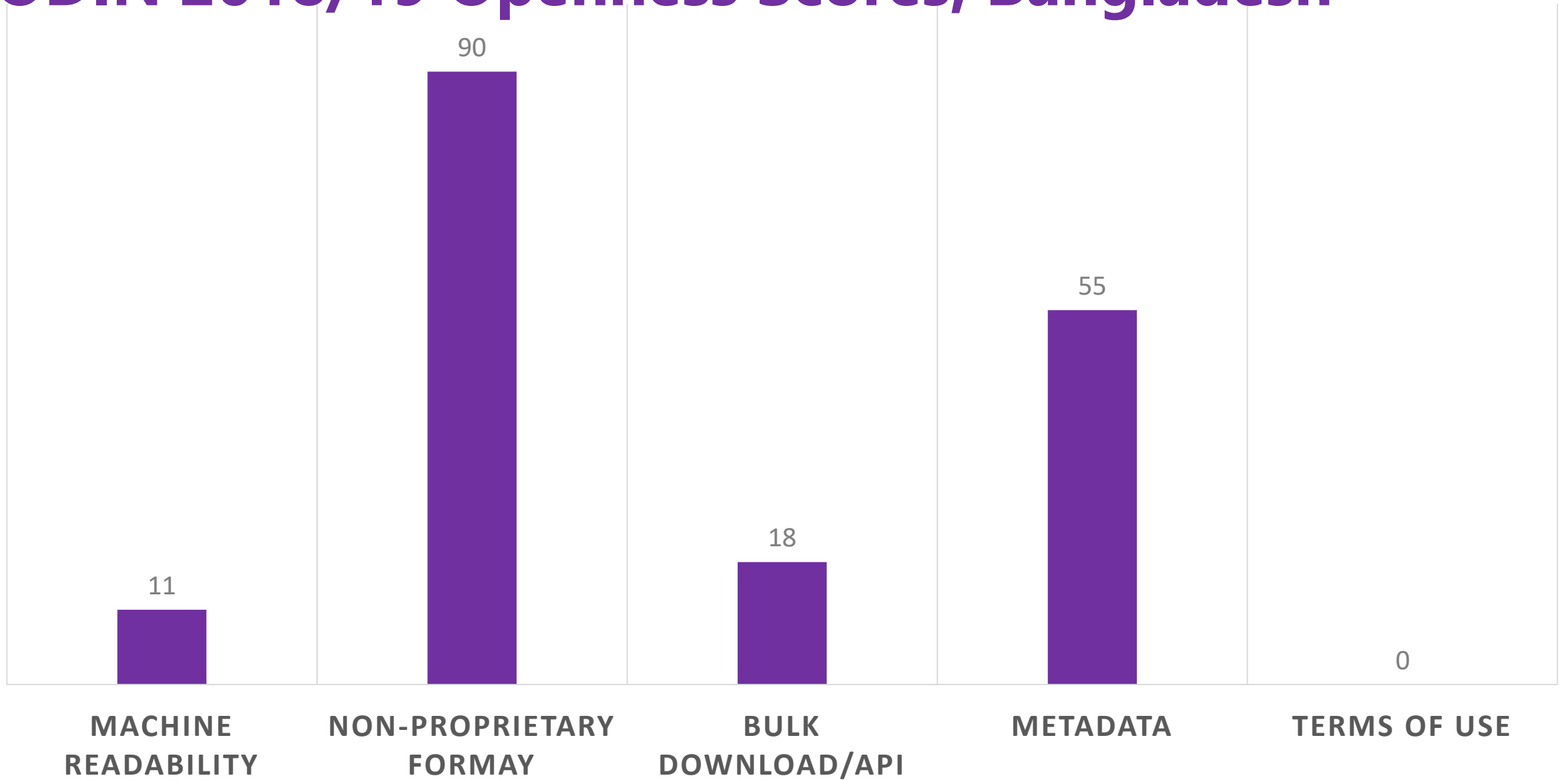
ODIN Scores by Category, Bangladesh



ODIN 2018/19 Coverage Scores, Bangladesh



ODIN 2018/19 Openness Scores, Bangladesh



Comparison of Indexes

Index	Which data?	Definition of openness	What does it measure?	How are assessments completed?
Open Data Inventory (ODIN)	National statistics posted on BBS website and linked agencies	Open Definition	Coverage and openness of key indicators	Expert researchers collect data. Government has a chance to review.
Open Data Barometer (ODB)	National government data	Open Definition	Coverage and openness of key datasets	Data is crowdsourced and reviewed. Public can comment.
Global Open Data Index (GODI)	National government data	Open Definition	Readiness, coverage and openness of key datasets, and impact	Expert researchers collect data. Secondary indicators used. Government does self-assessment.

Summary of Index Results

Index	Recommendations
Open Data Inventory (ODIN)	<ul style="list-style-type: none">• Publish data under a open data license• Publish data in machine readable formats• Make data available in bulk downloads• Publish more metadata• Publish more data at the subnational level
Open Data Barometer (ODB)	<ul style="list-style-type: none">• Publish data under a open data license• Publish data in machine readable and nonproprietary formats• Make data available in bulk downloads• Data needs to be updated more regularly• Publish more metadata
Global Open Data Index (GODI)	<ul style="list-style-type: none">• Publish data under a open data license• Publish data in machine readable and nonproprietary formats• Publish more data (data in few data categories were available)• Make data available in bulk downloads• Data needs to be updated more regularly



OPEN DATA LICENSING

PART 3:

International open data licenses, problematic clauses, and choosing a license

Schedule

Topics and activities



**Why does data
need a open license?**



Country Examples

Examples of international licenses,
country-created licenses, and
countries who've made
modifications.



**Group Exercise: Rewriting
your Terms of Use**

Try rewriting your own Terms of Use
so it is fully open

Why must data be licensed?

Adopting an open license is a core component of the definition of open data.

Data cannot be open unless they are licensed for reuse.

Adopting a license data is an opportunity for governments to encourage public use of their data by:

- specifically addressing how people can use data
- how they should attribute data
- alleviating user concerns about legal ramifications of unapproved use

What are the options?

International Standard Licenses

This includes licenses such as Creative Commons (CC), most commonly CC BY 4.0 and CC0.

Government-created Licenses

Various governments have also created their own open data licenses following the guidelines in the Open Definition.

Government-created Terms of Use

Governments have also created terms of use that include their data use policies.

Open Definition Guidelines on Licensing (Simple)

“Data that can be freely used, re-used and redistributed by anyone- subject only, at most, to the requirement to attribute and sharealike.”

Open Definition Guidelines on Licensing

2.1.1 Use

The license *must* allow **free use** of the licensed work.

2.1.2 Redistribution

The license *must* **allow redistribution** and sale

2.1.3 Modification

The license *must* allow the **creation of derivatives**

2.1.4 Separation

The license *must* allow any part of the work to be freely used, distributed, or modified **separately**

2.1.5 Compilation

The license *must* allow the licensed work to be **distributed along with** other distinct works

2.1.6 Non-discrimination

The license *must* **not discriminate** against any person or

group.

2.1.7 Propagation

The rights attached to the work *must apply to all to whom it is redistributed* without the need to agree to any additional legal terms.

2.1.8 Application to Any Purpose

The license *must* allow use, redistribution, modification, and compilation **for any purpose**. The license *must not* restrict anyone from making use of the work in a specific field of endeavor.

2.1.9 No Charge

Open Licenses around the World

Open License Status	% of Countries
All statistical data published under an open license	8%
No statistical data published under an open license	53%
Statistical data published under multiple licenses	39%

Creative Commons BY 4.0

You are free to:

Share — copy and redistribute the material in any medium or format

Adapt — remix, transform, and build upon the material
for any purpose, even commercially

Under the following terms:

Attribution — You must give appropriate credit, provide a link to the license, and indicate if changes were made.

No additional restrictions — You may not apply legal terms or technological measures that legally restrict others from doing anything the license permits.

Creative Commons 0

No Copyright

The person who associated a work with this deed has **dedicated** the work to the public domain by waiving all of his or her rights to the work worldwide under copyright law, including all related and neighboring rights, to the extent allowed by law.

You can copy, modify, distribute and perform the work, even for commercial purposes, all without asking permission. See **Other Information** below.

Example 1: Rwanda (CC BY 4.0)

”Hence, promoting widespread dissemination, use and sharing, data and analysis from the NISR is released under a [Creative Commons Attribution 4.0 International \(CC BY 4.0\)](#) license.

Accordingly, data and analysis from the NISR can be copied, distributed, transmitted and freely adapted, even for commercial purposes, provided that their integrity is respected and the source - “National Institute of Statistics of Rwanda” - is acknowledged appropriately.”

Example 2: Oman

“The data available in this site is open, and no license is required to use/re-use it.

When you come across to some statistics that may interfere privacy of others, kindly cooperate with the Centre to protect that privacy and inform the Centre about it.

The data should not be used illegally either where you are, or where we are.”

Example 3: Palestine

“User Rights

PCBS grants a universal, free-of-charge, irrevocable, parallel right of use to the material, Including:

- Copying, distributing, reusing, building, and deriving materials
- Editing and using for commercial or non-commercial purposes
- Using and quoting material in other publications

Note: These Terms of Use were drafted in guidance with the Creative Commons (CC) Attribution License.”

Problematic Clauses in Many Countries

Forbids Noncommercial Use

“No part or parts of this website may be modified, copied, distributed, retransmitted, broadcasted, displayed, reproduced, published, licensed, transferred, sold or commercially dealt with in any manner without the express prior written consent of the Department.” – Department of Statistics Malaysia

Forbids Misleading

“if reproduced, have to be accurate and are not to be used in a derogatory manner or in a misleading context” – Open Government Data Platform India

Problematic Clauses in Many Countries

Too Vague

“We wish to inform that our products are protected by copyright.

We expect all users to acknowledge Stats SL as the source of the basic data wherever they process, apply, utilise, publish or distribute the data, and also that they specify that the relevant application and analysis (where applicable) result from their own processing of the data.” – Statistics Ssierra Leone

Open Definition Guidelines on Licensing

2.1.1 Use

The license *must* allow **free use** of the licensed work.

2.1.2 Redistribution

The license *must* **allow redistribution** and sale

2.1.3 Modification

The license *must* allow the **creation of derivatives**

2.1.4 Separation

The license *must* allow any part of the work to be freely used, distributed, or modified **separately**

2.1.5 Compilation

The license *must* allow the licensed work to be **distributed along with** other distinct works

2.1.6 Non-discrimination

The license *must* **not discriminate** against any person or

group.

2.1.7 Propagation

The rights attached to the work *must apply to all to whom it is redistributed* without the need to agree to any additional legal terms.

2.1.8 Application to Any Purpose

The license *must* allow use, redistribution, modification, and compilation **for any purpose**. The license *must not* restrict anyone from making use of the work in a specific field of endeavor.

2.1.9 No Charge

Data.gov.bd Open Data Portal TOU

Bangladesh Government Web Portal aims to provide the citizens with a gateway to relevant services and updated information about Government services. This web portal is maintained by Prime Ministers' Office. By accessing and using this website people will be deemed to have accepted and be legally bound by these terms and conditions. Acceptance of the Terms and Conditions shall take effect from the time and date when the users first access this Website.

Disclaimer

The Prime Minister's Office does not sponsor, endorse or necessarily approve of any material on websites linked from or to this website;

This office accepts no liability for direct, indirect, special, punitive, incidental, exemplary or consequential damages arising out of accessing or using this website and websites linked from or to this website, or any use of the information and data or reliance placed on them; and

The Prime Minister's Office does not warrant that the functions contained in this Website will be uninterrupted or error free and that defect will be corrected;

Data.gov.bd Open Data Portal TOU, Page 2

Copying and Printing

Permission is granted to all user of this website to print all the information available in the websites without any deletion, addition or modification. This permission to print does not extend to any material on this website or any website to or from which this website is linked which is the copyright of any party other than the Government of Bangladesh.

Links to External Web Sites

This website contains links to other web sites, which are not maintained by the Prime Minister's Office or under its control are not responsible for the content of any such linked websites. The content and information of the websites linked to this portal will regularly be maintained and updated by the concerned organizations.

Data.gov.bd Open Data Portal TOU

Copying and Printing

Permission is granted to all user of this website to print all the information available in the websites **without any deletion, addition or modification**. This permission to print does not extend to any material on this website or any website to or from which this website is linked which is the copyright of any party other than the Government of Bangladesh.

Links to External Web Sites

This website contains links to other web sites, which are not maintained by the Prime Minister's Office or under its control are not responsible for the content of any such linked websites. The content and information of the websites linked to this portal will regularly be maintained and updated by the concerned organizations.

Data.gov.bd Open Data Portal TOU

Right of Access

The authority deserves all rights to deny or restrict access to this website to any particular person(s) or persons belonging to any geographical location or country or to block access from a particular Internet address at any time without assigning any reasons.

Policy Revision and Notification of Changes

The Prime Minister's Office of Bangladesh may at any time revise or update this policy without notice. Any information collected under the current policy will remain subject to these terms. Information collected after any changes take effect will be subject to the revised privacy policy.

Data.gov.bd Open Data Portal TOU

Right of Access

The authority deserves all rights **to deny or restrict access to this website to any particular person(s) or persons belonging to any geographical location or country or to block access from a particular Internet address at any time without assigning any reasons.**

Policy Revision and Notification of Changes

The Prime Minister's Office of Bangladesh **may at any time revise or update this policy without notice.** Any information collected under the current policy will remain subject to these terms. Information collected after any changes take effect will be subject to the revised privacy policy.

Data.gov.bd Open Data Portal TOU

Governing and Dispute Resolution

These Terms and Conditions shall be governed by and construed in accordance with the laws of Bangladesh. Any dispute arising under the Terms and Conditions shall be subject to the exclusive jurisdiction of the courts of Bangladesh. For more Information, please contact- Old Sangsad Bhaban, Tejgaon, Dhaka-1215 Bangladesh. E-mail: info@pmo.gov.bd



USER NEEDS & ENGAGEMENT

PART 4a:

Finding and addressing user needs

User-Centered Dissemination

The primary demand for data currently comes from international monitoring agencies, not local policy makers and citizens.

Therefore, greater attention should be placed on monitoring the use of data portals/websites and promoting them through targeted national user forums.

Identify Users

Before you bring publishers and data users together to address key challenges obtaining data, you must identify your users:

- public sector bodies
- government ministers and senior officials
- civil society organizations
- businesses
- local councils and local service providers
- journalists, politicians, educators, statisticians

Obtain Feedback

Methods	Pros	Cons
User surveys	<ul style="list-style-type: none">• Can be conducted in person or online• Inexpensive• Has a wide reach	<ul style="list-style-type: none">• Users have skip or quit questionnaire• If too long, fatigue could make responses inaccurate
Focus groups	<ul style="list-style-type: none">• More thorough and accurate responses• You know who your users are	<ul style="list-style-type: none">• More time consuming and costly• Participants may be affected by other's responses
Website Analytics	<ul style="list-style-type: none">• Relatively unbiased way to learn about user behavior• Many free options, such as Google Analytics	<ul style="list-style-type: none">• Requires some technical expertise to set up• Analytics doesn't tell the whole story

Ways to Make Data Easy to Find

1. Create a landing page categorized by topics, not report names
2. Ensure search functionality brings back expected results
3. If you redirect users, redirect them to specific pages where they can find data, not simply their home page
4. Create summaries below report links to let users know what they can expect to find inside



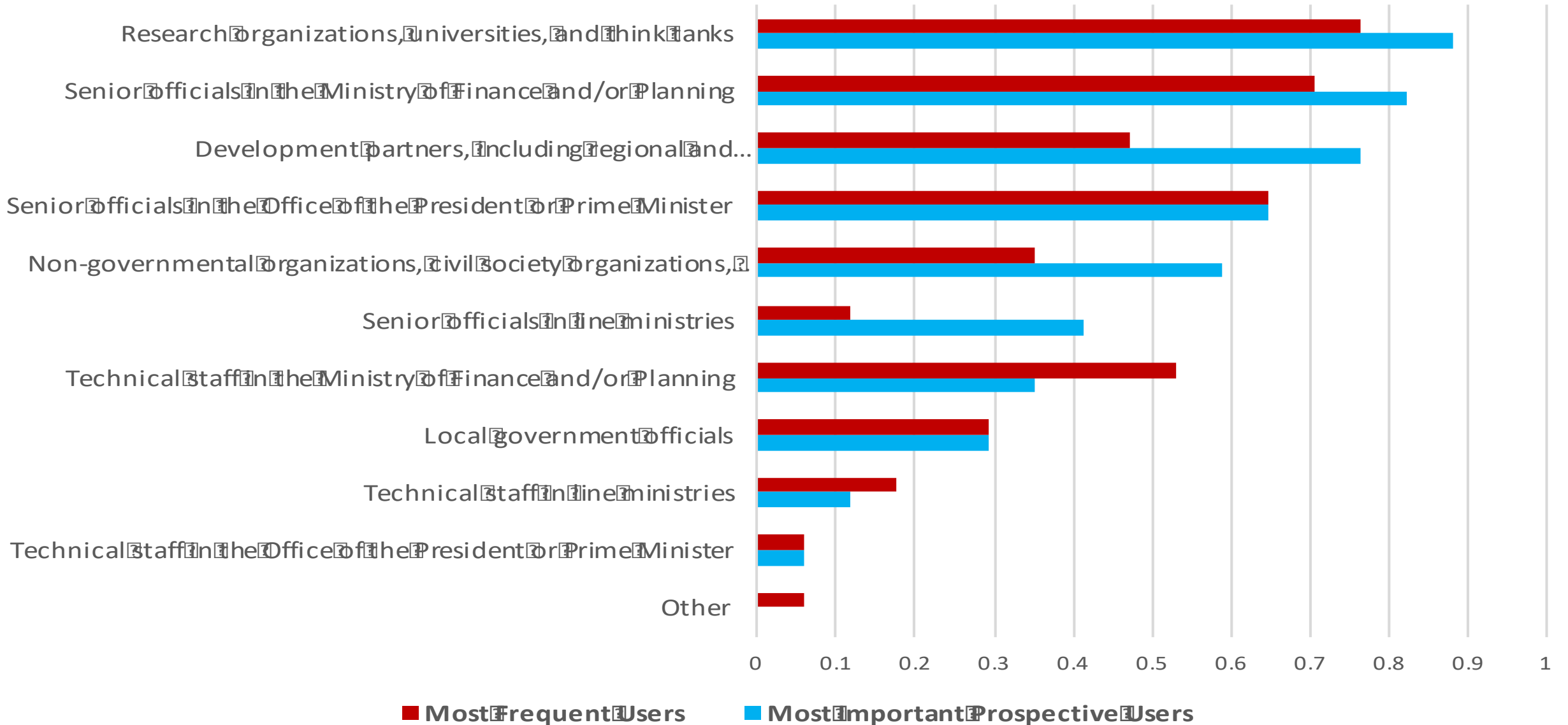
USER NEEDS SURVEY RESULTS

PART 4b: **Survey results and summary**

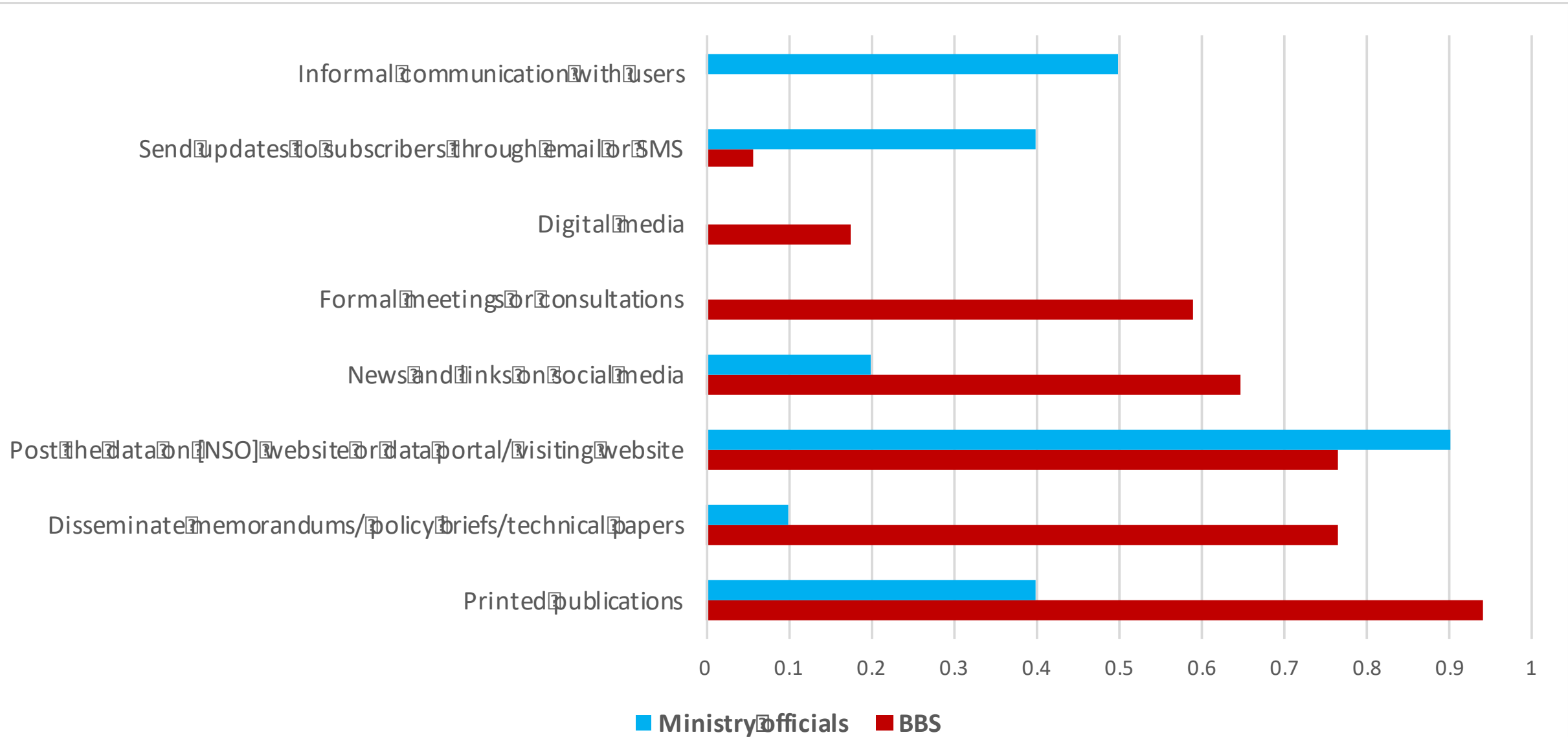
Type of Respondents

Ministry	Percentage
BBS	63%
Ministry of Health	7%
Central Bank	4%
Ministry of Local Gov	4%
Ministry of Education	4%
Ministry of Planning	7%
ICT Division	11%

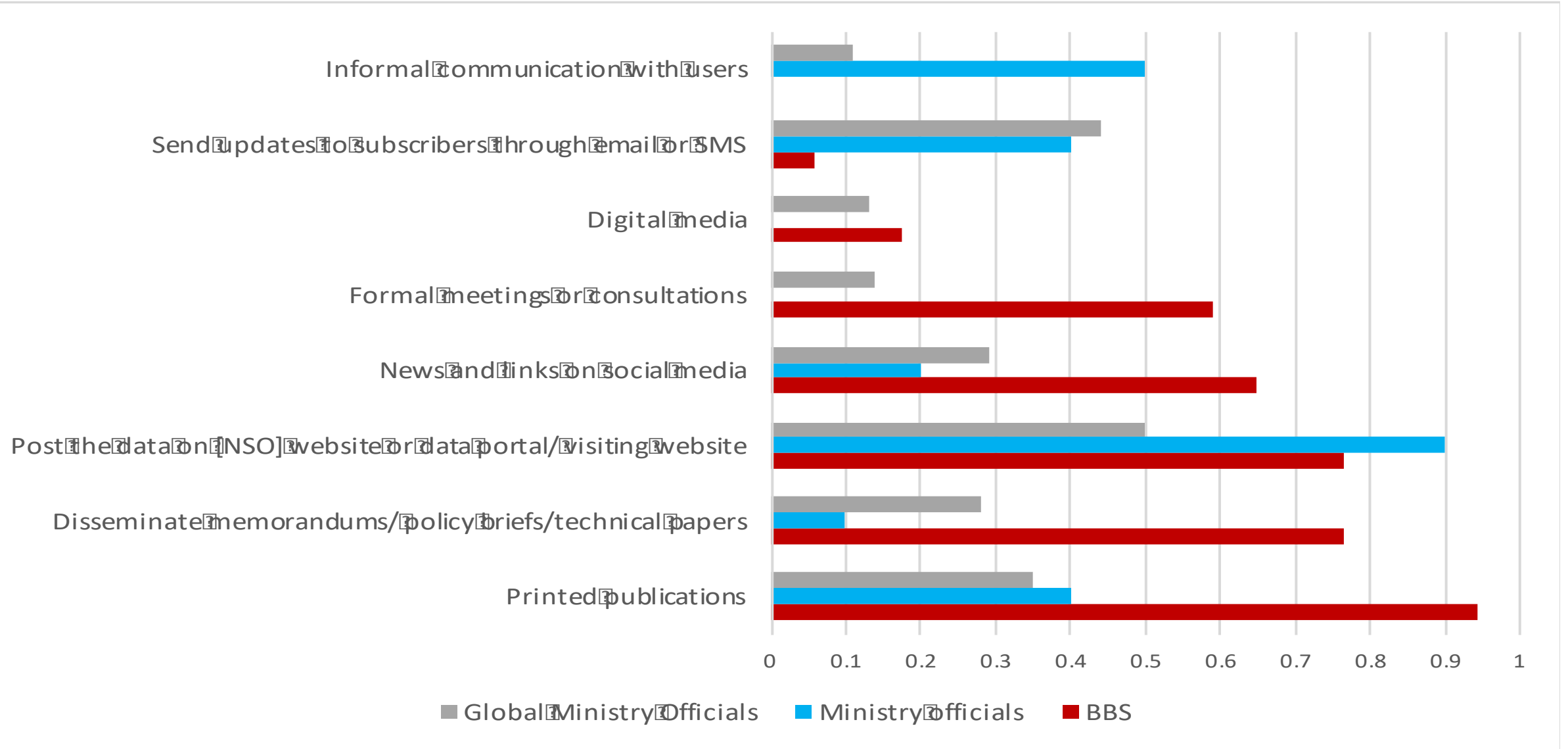
Who does BBS officials consider their most important and most frequent users?



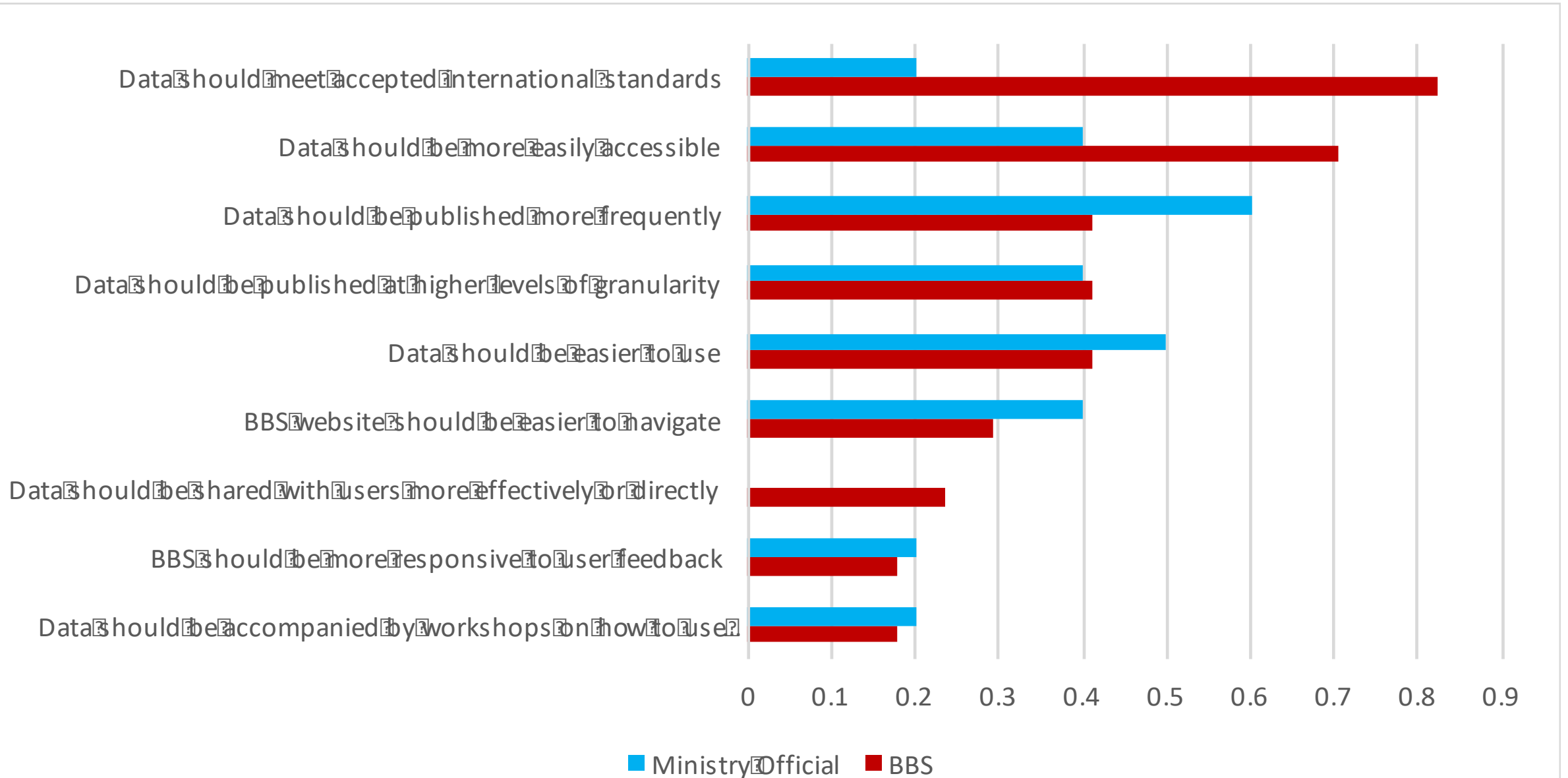
How does BBS inform their users about their data and how would ministry officials prefer to learn about these data?



How does BBS inform their users about their data and how would ministry officials prefer to learn about these data?



What improvements does BBS/Ministry Officials think are most important?



Does BBS measure the use of data it produces?

Answer	Percentage
Yes	6%
No	94%

How important is it for BBS to be able to measure the use of data it produces?

Answer	Percentage
Very important	59%
Quite important	35%
Not very important	6%

How would you like to measure data use?

