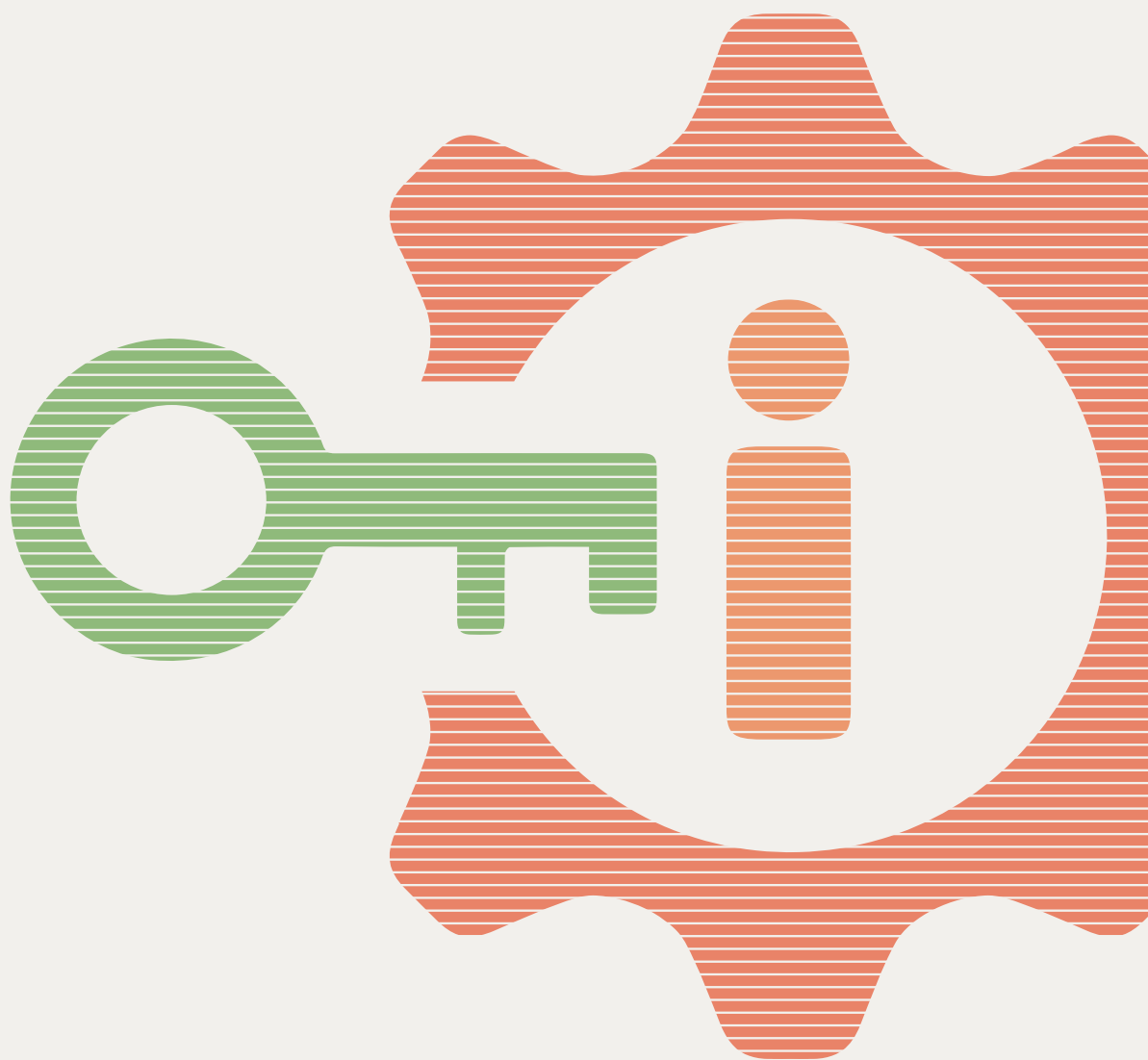




Development *and* Access to Information

2017



W
TECHNOLOGY &
SOCIAL CHANGE GROUP

UNIVERSITY of WASHINGTON
Information School

The International Federation of Library Associations and Institutions (IFLA) is the leading international body representing the interests of library and information services and their users. It is the global voice of the library and information profession.

The Technology & Social Change Group (TASCHA) at the University of Washington Information School explores the design, use, and effects of information and communication technologies in communities facing social and economic challenges. With experience in over 50 countries, TASCHA brings together a multidisciplinary network of researchers, practitioners, and policy experts to advance knowledge, create public resources, and improve policy and program design.

This report is funded by a grant from the Bill & Melinda Gates Foundation.

© 2017 by the International Federation of Library Associations and Institutions (IFLA) and the Technology and Social Change Group, University of Washington. This work is licensed under the Creative Commons Attribution 4.0 International (CC BY 4.0) license. To view a copy of this license, visit: <http://creativecommons.org/licenses/by/4.0>

IFLA
P.O. Box 95312
2509 CH Den Haag
Netherlands
www.ifla.org



Contact: DA2I@ifla.org
Website: <https://DA2I.ifla.org>

ISBN 978-90-77897-65-2 (Paperback)
ISBN 978-90-77897-67-6 (PDF)

ISSN 2588-9036 (Print)
ISSN 2588-9184 (Online)

Appendix 2: Data Curation, Processing, and Analysis Strategy

Using existing indicators for the baseline DA2I report and subsequent progress reports provides access to a vast amount of data on countries around the world, including economic and population data, and data on infrastructure and access to technology. The challenge is that when combining these indicators to build a more complete picture of the relationships between these dimensions, the variety of sources, data types, and collection strategies present some additional challenges beyond just the analysis. This section details the steps taken to compile the data, the processing done before any analysis, and finally the general data analysis strategy.

Data curation

To facilitate analysis and comparisons between the broad range of indicators selected for the baseline report, data for each indicator was compiled into a single database. This is an ongoing process that will continue for the life of the DA2I project, as new data becomes available. Indicator data was primarily sourced using databases from the International Telecommunication Union, World Bank, International Labour Organization, UN, UNESCO Institute of Statistics, Varieties of Democracy, and Freedom House. Although merging indicators from all these sources into a single database facilitated analysis and comparison between indicators from different official databases, multiple challenges to this approach were discovered during the process:

1. Different standards used for country names. To resolve this, country names as used by the World Bank were chosen as the standard, and data from all other sources was checked against this list, and any disparities fixed. This means that data extracted from the combined database for analysis has a consistent set of country names for all indicators.
2. Indicators are not consistently available for all countries for all years. In some cases, older data is available for fewer countries, but for others, only one data point is available for each country and the year data was collected varies. To mitigate this, as many years as possible were included in the combined database, so that if data was not available for a particular year, the most recent observation could be used instead. This also maximizes the flexibility as more indicators are added.
3. Each online database exports data in a slightly different format, requiring individual attention to ensure accurate import into the combined database. However, data extracted from the combined database is already in a consistent format, separating the data collection/import stage from the analysis stage.

Solving these challenges during the initial data curation phase simplified all of the subsequent analysis, reducing the possibility of errors in accessing and using the data.

Data processing

Two consistent challenges in preparing data subsets for analysis were due to inconsistencies in both the geographical and the time components of the indicator data. Although the data curation stage helped to mitigate these challenges, extra preparation of some indicators was required for analysis.

First, some indicators are only available for a subset of countries. For example, Freedom on the Net currently covers 65 countries, while the percentage of individuals using the internet by gender is available for 84 countries. The analytical problems are amplified in cases when two such indicators are compared, since the overlap between them in terms of which countries are covered by both indicators can be much smaller than either one individually. This impacted the types of analysis, most often preventing regional, income group, and world averages since not enough data was available for reliable estimates. Instead, in these cases analysis focused on presenting the data at the country level.

Secondly, the time components often required extra processing work when selecting data for a particular comparison between indicators. This is caused by two main factors: One, several indicators do not have a full panel of countries for each year; and two, some indicators have only one data point per country, collected in a variety of years. In these cases, a straight match using country and year between two indicators may return very few results. In some cases, one indicator in a comparison had a full panel for each year, which allowed matching by both country and year, effectively solving the problem. In other cases, however, the closest matching year for each country had to be computed for the two indicators, to compile the best available data

for each comparison. Hopefully this challenge will be mitigated in the future as more data is collected, but it is likely to remain a problem at some level.

Data analysis strategy

In general, analysis proceeded from a global view, to views by income group (using the World Bank categories) and region (from the UN Sustainable Development Goals report classification), and finally down to a country level. Disaggregating by income group and region bridges the gap between global averages and individual countries, offering a useful lens through which to view the data. Since countries can have vastly different population sizes, indicators that measured normalized values (percentages or counts per 100 inhabitants) were weighted by country population in world, income group, and regional averages, in order to make aggregate values more representative.



UN Sustainable Development Goal Region List

Country	Region Classification
Armenia	Caucasus and Central Asia
Azerbaijan	Caucasus and Central Asia
Georgia	Caucasus and Central Asia
Kazakhstan	Caucasus and Central Asia
Kyrgyz Republic	Caucasus and Central Asia
Tajikistan	Caucasus and Central Asia
Turkmenistan	Caucasus and Central Asia
Uzbekistan	Caucasus and Central Asia
Albania	Developed regions
Andorra	Developed regions
Australia	Developed regions
Austria	Developed regions
Belarus	Developed regions
Belgium	Developed regions
Bermuda	Developed regions
Bosnia and Herzegovina	Developed regions
Bulgaria	Developed regions
Canada	Developed regions
Channel Islands	Developed regions
Croatia	Developed regions
Cyprus	Developed regions
Czech Republic	Developed regions
Denmark	Developed regions
Estonia	Developed regions
Faroe Islands	Developed regions
Finland	Developed regions
France	Developed regions
Germany	Developed regions
Greece	Developed regions

Greenland	Developed regions
Hungary	Developed regions
Iceland	Developed regions
Ireland	Developed regions
Isle of Man	Developed regions
Israel	Developed regions
Italy	Developed regions
Japan	Developed regions
Latvia	Developed regions
Liechtenstein	Developed regions
Lithuania	Developed regions
Luxembourg	Developed regions
Macedonia, FYR	Developed regions
Malta	Developed regions
Moldova	Developed regions
Monaco	Developed regions
Montenegro	Developed regions
Netherlands	Developed regions
New Zealand	Developed regions
Norway	Developed regions
Poland	Developed regions
Portugal	Developed regions
Romania	Developed regions
Russian Federation	Developed regions
San Marino	Developed regions
Serbia	Developed regions
Slovak Republic	Developed regions
Slovenia	Developed regions
Spain	Developed regions
Sweden	Developed regions
Switzerland	Developed regions
Ukraine	Developed regions
United Kingdom	Developed regions
United States	Developed regions
China	Eastern Asia
Hong Kong SAR, China	Eastern Asia
Korea, Dem. People's Rep.	Eastern Asia
Korea, Rep.	Eastern Asia
Macao SAR, China	Eastern Asia
Mongolia	Eastern Asia
Anguilla	Latin America and the Caribbean
Antigua and Barbuda	Latin America and the Caribbean
Argentina	Latin America and the Caribbean
Aruba	Latin America and the Caribbean
Bahamas, The	Latin America and the Caribbean
Barbados	Latin America and the Caribbean
Belize	Latin America and the Caribbean
Bolivia	Latin America and the Caribbean
Bonaire, Sint Eustatius and Saba	Latin America and the Caribbean
Brazil	Latin America and the Caribbean

British Virgin Islands	Latin America and the Caribbean
Cayman Islands	Latin America and the Caribbean
Chile	Latin America and the Caribbean
Colombia	Latin America and the Caribbean
Costa Rica	Latin America and the Caribbean
Cuba	Latin America and the Caribbean
Curaçao	Latin America and the Caribbean
Dominica	Latin America and the Caribbean
Dominican Republic	Latin America and the Caribbean
Ecuador	Latin America and the Caribbean
El Salvador	Latin America and the Caribbean
Falkland Islands (Malvinas)	Latin America and the Caribbean
French Guiana	Latin America and the Caribbean
Grenada	Latin America and the Caribbean
Guadeloupe	Latin America and the Caribbean
Guatemala	Latin America and the Caribbean
Guyana	Latin America and the Caribbean
Haiti	Latin America and the Caribbean
Honduras	Latin America and the Caribbean
Jamaica	Latin America and the Caribbean
Martinique	Latin America and the Caribbean
Mexico	Latin America and the Caribbean
Montserrat	Latin America and the Caribbean
Nicaragua	Latin America and the Caribbean
Panama	Latin America and the Caribbean
Paraguay	Latin America and the Caribbean
Peru	Latin America and the Caribbean
Puerto Rico	Latin America and the Caribbean
Sint Maarten (Dutch part)	Latin America and the Caribbean
St. Kitts and Nevis	Latin America and the Caribbean
St. Lucia	Latin America and the Caribbean
St. Vincent and the Grenadines	Latin America and the Caribbean
Suriname	Latin America and the Caribbean
Trinidad and Tobago	Latin America and the Caribbean
Turks and Caicos Islands	Latin America and the Caribbean
Uruguay	Latin America and the Caribbean
Venezuela, RB	Latin America and the Caribbean
Virgin Islands (U.S.)	Latin America and the Caribbean
Algeria	Northern Africa
Egypt, Arab Rep.	Northern Africa
Libya	Northern Africa
Morocco	Northern Africa
Tunisia	Northern Africa
Western Sahara	Northern Africa
American Samoa	Oceania
Cook Islands	Oceania
Fiji	Oceania
French Polynesia	Oceania
Guam	Oceania
Kiribati	Oceania

Marshall Islands	Oceania
Micronesia, Fed. Sts.	Oceania
Nauru	Oceania
New Caledonia	Oceania
Niue	Oceania
Northern Mariana Islands	Oceania
Palau	Oceania
Papua New Guinea	Oceania
Samoa	Oceania
Solomon Islands	Oceania
Tokelau	Oceania
Tonga	Oceania
Tuvalu	Oceania
Vanuatu	Oceania
Brunei Darussalam	South-eastern Asia
Cambodia	South-eastern Asia
Indonesia	South-eastern Asia
Lao PDR	South-eastern Asia
Malaysia	South-eastern Asia
Myanmar	South-eastern Asia
Philippines	South-eastern Asia
Singapore	South-eastern Asia
Thailand	South-eastern Asia
Timor-Leste	South-eastern Asia
Vietnam	South-eastern Asia
Afghanistan	Southern Asia
Bangladesh	Southern Asia
Bhutan	Southern Asia
India	Southern Asia
Iran, Islamic Rep.	Southern Asia
Maldives	Southern Asia
Nepal	Southern Asia
Pakistan	Southern Asia
Sri Lanka	Southern Asia
Angola	Sub-Saharan Africa
Benin	Sub-Saharan Africa
Botswana	Sub-Saharan Africa
Burkina Faso	Sub-Saharan Africa
Burundi	Sub-Saharan Africa
Côte d'Ivoire	Sub-Saharan Africa
Cabo Verde	Sub-Saharan Africa
Cameroon	Sub-Saharan Africa
Central African Republic	Sub-Saharan Africa
Chad	Sub-Saharan Africa
Comoros	Sub-Saharan Africa
Congo	Sub-Saharan Africa
Congo, Dem. Rep.	Sub-Saharan Africa
Djibouti	Sub-Saharan Africa
Equatorial Guinea	Sub-Saharan Africa
Eritrea	Sub-Saharan Africa
Ethiopia	Sub-Saharan Africa

Gabon	Sub-Saharan Africa
Gambia, The	Sub-Saharan Africa
Ghana	Sub-Saharan Africa
Guinea	Sub-Saharan Africa
Guinea-Bissau	Sub-Saharan Africa
Kenya	Sub-Saharan Africa
Lesotho	Sub-Saharan Africa
Liberia	Sub-Saharan Africa
Madagascar	Sub-Saharan Africa
Malawi	Sub-Saharan Africa
Mali	Sub-Saharan Africa
Mauritania	Sub-Saharan Africa
Mauritius	Sub-Saharan Africa
Mayotte	Sub-Saharan Africa
Mozambique	Sub-Saharan Africa
Namibia	Sub-Saharan Africa
Niger	Sub-Saharan Africa
Nigeria	Sub-Saharan Africa
Réunion	Sub-Saharan Africa
Rwanda	Sub-Saharan Africa
Sao Tome and Principe	Sub-Saharan Africa
Senegal	Sub-Saharan Africa
Seychelles	Sub-Saharan Africa
Sierra Leone	Sub-Saharan Africa
Somalia	Sub-Saharan Africa
South Africa	Sub-Saharan Africa
South Sudan	Sub-Saharan Africa
Sudan	Sub-Saharan Africa
Swaziland	Sub-Saharan Africa
Tanzania	Sub-Saharan Africa
Togo	Sub-Saharan Africa
Uganda	Sub-Saharan Africa
Zambia	Sub-Saharan Africa
Zimbabwe	Sub-Saharan Africa
Bahrain	Western Asia
Iraq	Western Asia
Jordan	Western Asia
Kuwait	Western Asia
Lebanon	Western Asia
Oman	Western Asia
Qatar	Western Asia
Saudi Arabia	Western Asia
State of Palestine	Western Asia
Syrian Arab Republic	Western Asia
Turkey	Western Asia
United Arab Emirates	Western Asia
Yemen, Rep.	Western Asia

