

Understanding Usability of SDG National Reporting Platforms



By Yu-En Hsu and Elettra Baldi

November 2020

After the adoption of the Sustainable Development Goals (SDGs) in 2015, countries began incorporating the SDGs into national monitoring and evaluation frameworks, creating SDG reporting platforms or revamping existing data portals to publish and share SDG data. These efforts are an essential component of target 17.18.1, which focuses on improving statistical capacity for monitoring the SDGs, and an important step towards achieving the 2030 Agenda.

National SDG reporting platforms are designed to store, manage, and disseminate SDG data and metadata. They allow citizens, policymakers, and other stakeholders to easily find, view, and download data for analysis and decision-making. But the technical infrastructure and user experience of national reporting platforms, if not designed properly, can stifle data use.

In May 2019, Open Data Watch (ODW) and **Center for Open Data Enterprise (CODE)** published a **blog** on the general development of SDG portals and potential areas for improvement. Over the past year, ODW has monitored several national SDG reporting platforms and expanded our review to include issues from users' perspectives to further our understanding of data use. This post discusses the main problems ODW encountered when trying to access, search, and locate the contents of national SDG reporting platforms. Each issue is illustrated with a strong and a problematic practice and recommendations drawn from our experience.

Can users access the national SDG reporting platform?

The first step in accessing a country's SDG data is to visit the national SDG reporting platform where the data are hosted and search for the desired datasets. The most common issues encountered are website downtime and slow loading speed. For datasets hyperlinked by portals, broken links are an additional concern. Both downtime, slow loading, and broken links discourage visitors from using or returning to the portal.

Website availability and speed

Website availability (also called website uptime) refers to the ability of the users to access and use a website or web service,¹ and website speed refers to how quickly a browser can load fully functional webpages from a given site.² These are critical factors that affect the use of national SDG reporting platforms. Consider these examples:

- **Strong practice:** Portal A maintains 99.9% website uptime with fast loading speed.
- **Problematic practice:** Portal B is inaccessible from time to time with a 503 Service Unavailable response code, indicating issues on the server site. When the portal is online, each webpage takes a long time to load. Sometimes the webpage goes offline after loading for an extended period.

¹ Uptrends, 2020, "What is website availability?" Uptrends Glossary. <https://www.uptrends.com/what-is/website-availability>

² Cloudflare, 2020, "Why Does Site Speed Matter?" Cloudflare. <https://www.cloudflare.com/learning/performance/why-site-speed-matters/>

- **Recommendation:** Provide a consistently available and fast portal allows users from different settings and time zones to access data. Poor performance is likely caused by the infrastructure or hardware on the server site; however, it may also be a problem of poorly implemented software. Regardless, national statistical offices (NSOs) and webmasters can utilize website monitoring tools, such as [Google PageSpeed](#), to establish a baseline, identify problems, and make improvements.

Broken links

Broken links on a website frustrate users trying to access the content and can reduce a website's ranking in search engines. For the purpose of the analysis the term "broken links" is used in a broad sense to refer to dead links that do not work or links that lead to an error or a wrong page. Ensuring all links are working and directing users to correct webpages provides a smooth user experience and easy navigation.

- **Strong practice:** Portal A allows a seamless transition among webpages. All the links are correctly set up and lead users to the intended destinations.
- **Problematic practice:** Portal B displays many links on the homepage, that lead to an error page. Users may need to find the destination webpage via Google Search or reach out to the portal manager to obtain data. During the monitoring, ODW researchers noticed obvious mistakes in hyperlinks, such as a comma at the end of the URL. ODW researchers were able to access the webpage after removing the comma. Due to human error, these small mistakes made it difficult for the user to access the data portal.
- **Recommendation:** Examine hyperlinks before publishing new webpages and monitor them afterward. When websites use HTTPS in the domain rather than HTTP, portal managers should make sure the Transport Layer Security (TLS) or its predecessor, the Secure Sockets Layer (SSL), is regularly renewed.

Can users find the data easily?

National SDG reporting platforms contain many datasets. Their interface should allow users to locate desired information easily and quickly.

Search function

Providing a search function offers users a way to find content by searching for words or phrases, without needing to understand or navigate the structure of the website.³ The ODW team often searches SDG indicator keywords to find relevant datasets. Take SDG indicator 9.2.2 (manufacturing employment as a proportion of total employment) as an example. To find this indicator a user might begin by searching for "manufacturing."

- **Strong practice:** Portal A's search function finds keywords from across the entire portal, including dataset titles, metadata, publications, and other materials. For example, a search for "manufacturing" returns "manufacturing wage index" and "working population by sector." The first result has the keyword in the dataset title, and the second one does not. But the "working population by sector" dataset contains the keyword in the industry group selection and has data for manufacturing employment. Moreover, the result page allows users to filter the type of result, such as statistics, announcements, and metadata.
- **Problematic practice:** Portal B's search function only looks for the keyword in the dataset title. In this case, the result page shows "manufacturing wage index" and "gross added value"

³ World Wide Web Consortium (W3C). (n.d.). G161: Providing a search function to help users find content | Techniques for WCAG 2.0. Retrieved 29 July 2020, from <https://www.w3.org/TR/WCAG20-TECHS/G161.html>

for manufacturing,” neither of which contain data on manufacturing employment. Data on manufacturing employment exist, but because the dataset is titled “employees by economic activity, sex and age,” the keyword “manufacturing” does not return a hit.

- **Recommendation:** Provide a search function that covers all components of the portal, such as dataset titles, dimensions, reports, and more. Additionally, the search function should not be case sensitive and should return approximate results. For example, when users search “mobilisation,” the function should look for “mobilization” as well, and when users search “disability,” the result should include “disabilities.” The search function needs to be comprehensive yet flexible. To ensure that datasets can be found, NSOs should include a robust set of keywords in the description of the dataset and in the metadata, as a search function can only be effective if it has a rich set of data descriptions, titles, and keywords to draw from. More advanced statistical offices may consider implementing aspects of semantic search into their website search systems, to provide more intuitive search results that are not based only on keyword matches.

Data interoperability

Interoperability is the ability to join-up and merge data without losing meaning.⁴ In practice, data are said to be interoperable when they can be easily re-used and processed in different applications, allowing different information systems to work together.⁵ For the purpose of our research, there are two ways of interpreting interoperability. First, SDG reporting platforms should assist users by merging dataset into continuous series. ODW researchers have noticed that many portals provide data only from 2015 onward, the year when SDGs were implemented, even though many data were collected before 2015. Second, the data should be interoperable with other datasets. Below, SDG indicator 3.1.1, maternal mortality ratio, is used as an example.

- **Strong practice:** Portal A provides maternal mortality statistics from 2000 to 2019. Although the Ministry of Health was responsible for compiling data from 2000 to 2015 and the NSO took over in 2015, portal A has all information in one place. Moreover, the dataset includes raw data on the number of deaths, the number of births, and International Classification of Diseases (ICD) identification code. The ICD code makes it easy for users to find the same classification in other portals.
- **Problematic practice:** Portal B, which is designed as an SDG reporting platform, has data from 2015 to 2018 only, even though data for earlier years are available in the NSO portal. In this case, users have to navigate multiple sites to gather complete historical data. Moreover, two portals use different units, one with the number of deaths and one with the mortality ratio, making it difficult and time-consuming for users to construct a continuous series.
- **Recommendation:** Connect different sources and provide all relevant data. For SDG reporting, the platforms should include all available data from before the SDGs initiation in 2015. SDG indicators should be presented in the units specified by their definition, but the portal should also provide raw data from which ratios, shares, or growth rates can be computed. For indicators that conform to international classifications, such as the International Standard Classification of Occupations (ISCO) or International Classification of Goods and Services (Nice), their classification codes should be included. It is understandable that some countries have

⁴ Liz Steele and Tom Orrell, 2017, “The frontiers of data interoperability for sustainable development,” Development Initiatives and Publish What You Fund. http://www.publishwhatyoufund.org/wp-content/uploads/2017/11/JUDS_Report_Web_061117.pdf

⁵ Luis González Morales and Tom Orrell, 2018, “Data Interoperability: A Practitioner’s Guide to Joining up Data in the Development Sector.” Global Partnership for Effective Development Co-operation. <https://www.effectivecooperation.org/content/data-interoperability-practitioners-guide-joining-data-development-sector>

adopted different systems, and in this case, the portal should provide information on the classification and a crosswalk to the international system.

Can users understand and use the data?

Once users find the desired dataset, the next steps include generating, browsing, visualizing, and exporting data. And they may still need to know more about the methodology and sources of the data.

Metadata availability and quality

Metadata, data about data, describe the how the data were constructed and define all pertinent terms. This is particularly useful when users find potential errors in data or have questions about the data's applicability. According to ODW's [Open Data Inventory \(ODIN\)](#), the metadata should at a minimum provide a definition of the indicator or definitions of key terms used in the indicator description and a description of how the indicator was calculated; the publication date (date of upload), compilation date, or the date of the last update; and the name of the data source. Take SDG indicator 4.1.2 (Completion rate (primary education, lower secondary education, upper secondary education)), as an example.

- **Strong practice:** Portal A displays metadata on the dataset page, and the metadata are complete, including data source, definition, tabulation, and publication date. The document also explains the education system in the country with expected entering age and required years. The metadata can be downloaded with the indicator.
- **Problematic practice:** Portal B has incomplete metadata with only data source specified. Without knowing how the indicator was calculated and definition for each education level, users may incorrectly interpret the data.
- **Recommendation:** Provide complete metadata that enables users to understand the indicator, the definition, and the source of the data.

Machine-readable download options

Machine-readable file formats, such as XLS, XLSX, CSV, or JSON, allow users to easily process data using a computer. When data are made available in formats that are not machine readable, users cannot easily access and modify the data, severely restricting the scope of data use.

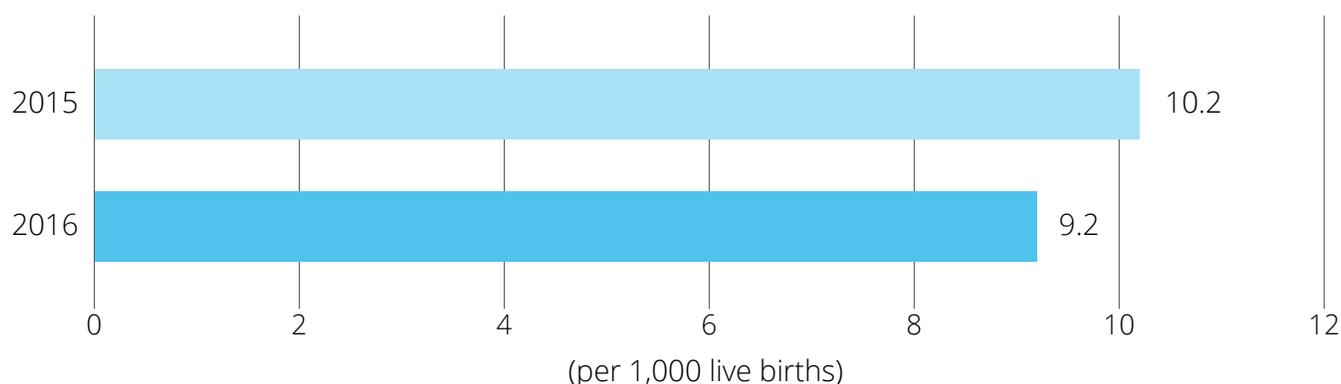
- **Strong practice:** Portal A provides multiple machine-readable download options: JSON, CSV, and XLSX. It also has an application programming interface (API) for more advanced users.
- **Problematic practice:** Portal B displays an HTML table but provides only PDF and PNG options for downloading. While users can still extract data manually, the process is time-consuming.
- **Recommendation:** Provide at least one machine-readable download option. For portals that already publish machine-readable data, considering providing additional formats and bulk download options.

Visualization clarity

Visualizations increase understanding and interpretation of data. Bar charts and line charts are the most common types offered in portals. Good visualizations can communicate dense information and allow users to observe patterns easily. Bad visualizations may distort data and can be misleading. Currently, there are no established standards or guidelines, but [Material](#), a design system by Google, provides some conventional methods and criteria for selecting and creating visualizations. The following uses SDG indicator 3.2.2 (Neonatal mortality rate) as an example. Both visualizations display values of 10.2 in 2015 and 9.2 in 2016.

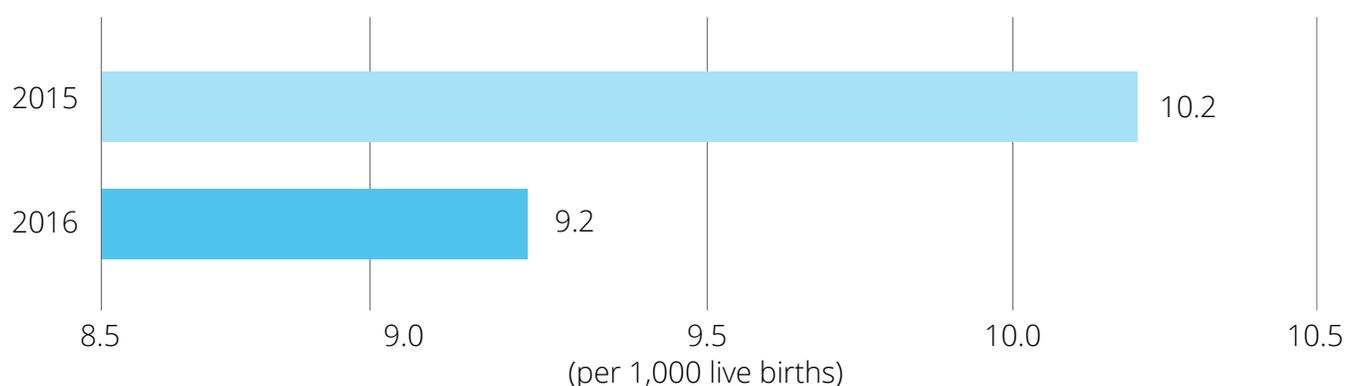
- **Strong practice:** Visualization A's x-axis starts from zero and ends at 12. The bar chart shows a slight decrease in mortality rates from 2015 to 2016.

Visualization A - Neonatal Mortality Rate



- **Problematic practice:** Visualization B's x-axis starts from 8.5 and ends at 10.5. At first glance, the visualization suggests the mortality in 2016 is less than half of that in 2015.

Visualization B - Neonatal Mortality Rate



- **Recommendation:** Review all visualizations for clarity and consistent representation of statistical relationships.

Identifying missing data

Portals should have a distinct label for data points that are unavailable, or they should be left blank, instead of inserting a value of 0. The most common identifier for no data is "N/A," an abbreviation for not available. Indicator 16.10.1 (Number of verified cases of killing, kidnapping, enforced disappearance, arbitrary detention and torture of journalists, associated media personnel, trade unionists and human rights advocates in the previous 12 months) is an example.

- **Strong practice:** Portal A shows 1 for 2015, 0 for 2016, and "N/A" for 2017. The zero in 2016 is an actual value. The distinctive label for 2017 confirms no data are available.
- **Problematic practice:** Portal B shows 1 in 2015, 0 in 2016, and 0 in 2017. This appears to show no verified cases of killing in 2016 and 2017. However, in this case, metadata state that the data are available up to 2016, suggesting the 0 for 2017 is referring to no data.

- **Recommendation:** Label data points that have no data with “N/A” or another unique label instead of zero, to ensure a clear understanding of missing data points.

Takeaways

A critical review of the design and performance of a data portal is important for ensuring that data are available and can be used. SDG reporting platforms should make an extra effort to accommodate a wide variety of users, some of whom may have little experience with statistical data, while others are constructing sophisticated applications that depend on reliable access to high-quality data. NSOs should also provide space on their SDG reporting platform for user feedback to enhance their portal based on commentary from their primary audience.

Open Data Watch encourages NSOs to conduct an internal evaluation of their portals from a user perspective and to adopt guidelines, such as the [Principles of SDG Indicator Reporting and Dissemination Platforms and guidelines for their application](#),⁶ to guide their development. For additional support, NSO's can consult ODW's [Data Site Evaluation Toolkit \(DSET\)](#) to assess and improve their national SDG reporting platforms. Our most recent DSET report, written with the World Bank Nepal Office, [Assessing the Effectiveness of Data Sites in Nepal](#), provides a demonstration of the toolkit. Taking these steps will help to ensure that data are not only made available on a portal but can be used to achieve the promises of the SDGs.

⁶ United Nations Statistics Division / Department of Economic and Social Affairs. (2019) Principles of SDG Indicator Reporting and Dissemination Platforms and guidelines for their application. <https://unstats.un.org/unsd/statcom/50th-session/documents/BG-Item3a-Principles-guidelines-SDG-Monitoring-Reporting-Platforms-E.pdf>