ASSESSING THE EFFECTIVENESS OF DATA SITES IN NEPAL

March 2020

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Table of Contents

ACKNOWLEDGMENTS

DISCLAIMER

INTRODUCTION

In the last few years a group of private and nonprofit organizations have been working to increase the public access to and use of data in Nepal. As part of this effort, organizations and government agencies, often with the support from development partners, have created publicly available data portals to provide access to data in the country.

While the creation of publicly available data portals are steps in the right direction, they must be findable, accessible, and usable by the general public. This study assesses the functionality of the selected data portals in Nepal by using the Data Site Evaluation Toolkit (DSET). The DSET is an evaluation system developed by Open Data Watch (ODW) to identify the elements of a well-functioning and open data site. The three main questions driving the DSET evaluation are:

CAN SOMEONE FIND THE SITE?

This section analyzes elements of search engine optimization and dissemination strategies to measure how the site appears in search engines and other findability issues.

HOW ACCESSIBLE IS THE SITE?

This section analyzes accessibility in different browsers and across devices and evaluates website speed, navigation, language accessibility, and user support options.

CAN SOMEONE USE THE DATA ON THE SITE?

This section utilizes Open Data Inventory (ODIN) criteria for openness and the availability of metadata and analyzes the availability of visualizations and usability issues that come up when testing the site.

This study evaluated 22 publicly available data portals that features data related to Nepal. This is a purposive and non-probability sample, primarily because there is no exhaustive list of data portals in Nepal that can reliably serve as a sampling frame to draw a random sample. As such, findings in this study are likely to be indicative of common issues in Nepal but cannot be generalized beyond the study sample. While the sample size is determined based on various factors including time and financial resources available at the time of the analysis,[1] the focus is on data portals that uses

---

[1] This study was conducted as one of the background papers for World Bank (2019) published in December 2019.

government data; the sample includes 9 portals maintained by government agencies and 13 portals that make use of the data produced by Government of Nepal.

Open Data Watch evaluated the selected 22 data sites against 71 elements of data portal functionality from DSET. Assessments of the sites were carried out between 8 October and 18 November 2019. Two of the sites were either not online (National Geospatial Portal) throughout the evaluation or failed to load their portal (Hydropower Projects in Nepal). These sites were only partially evaluated. Complete evaluations were conducted for 20 sites.

Further, the DSET evaluation is focused on the elements of a well-functioning and open data site and not on availability and reliability of datasets on those sites, although the research team acknowledges that this is an important factor for data use. Other existing studies, including the Open Data Inventory and Open Data for Resilience Initiative, can provide more information on data gaps in Nepal.

The results of this study provide insights into the common issues a user might encounter when accessing, finding, and using data in Nepal. It also provides actionable recommendations on how development partners, portal sponsors and creators, and data portal managers can improve their sites to meet the needs of these users. This report is intended for stakeholders interested in making data in Nepal more findable and usable. The audience includes representatives of the government, development partners, academia, civil society, the media, and other groups.

EVALUATION RESULTS

CAN SOMEONE FIND THE SITE?

The majority of sites studied are not using best practices for search engine optimization (SEO) and do not appear in many search results. SEO is a process by which sites can organically boost the quantity and quality of their traffic via search results on search engines such as Google. Consequently, many of these sites may not receive much traffic. The sites that perform better are linked with more established domains that have invested in best practices for SEO such as sitemaps and blogs. The sites that performed poorly on this portion of the DSET evaluation have few external sites linking to them. When a site is linked by another site, it is a sign of trust and a crucial metric for SEO. External links are an important metric that can increase a site's place in search engine rankings and are an important source of traffic for sites.

Missing SEO and findability elements

Organic search traffic is typically one of the biggest sources of visitors to a website. If a site is not implementing best practices for SEO, Google's spiders may not be able to find it or face issues crawling the site, which can hurt its rankings in search engine results. Many of the sites didn't have an XML site map, a critical SEO tool, and received low scores on metrics from Moz, an SEO company, that predict rankings in search engines and take into account dozens of SEO metrics. Further, only seven of the listed sites had installed Google Analytics, a tool that can be used to track traffic to a site and be leveraged to bring in more users.

Figure 1: Findings from the SEO and findability assessment
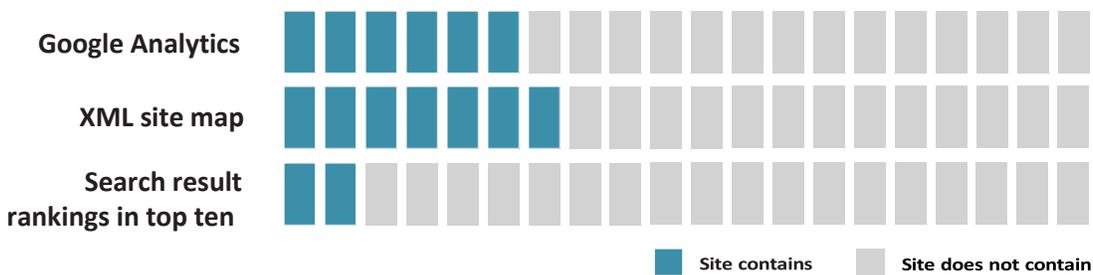
Low uptime for sites

Website uptime (time a website is online) is a critical measurement for SEO and a precondition for a user to access a site. A number of websites experienced significant downtime during the evaluation. There is not a clear industry standard for uptime of a website. High availability – or an uptime percentage of 99 percent or more – is one measure. This may sound like a high standard, but even a site with 99 percent uptime is down for 87 hours and 36 minutes a year.

Three sites in the evaluation had uptime scores short of the 99 percent standard. Although the Hydropower Projects in Nepal site failed to load the portal, the site's uptime was 99 percent and the server was responding so it was not counted as having low uptime in this exercise.

Table 1: Sites that experienced significant downtime during the evaluation

Website Name	URL	Time Online (%)
Housing Recovery and Reconstruction Platform	hrrpnepal.org	97.0
National Data Portal-Nepal	nationaldata.gov.np	51.0
National Geospatial Portal	nationalgeoportal.gov.np	0.0

Regional uptime comparisons

Website uptime in the region provides context for how the sites studied in Nepal compare to sites in countries surrounding Nepal. ODW analyzed the uptime for NSO websites around the world from July 2018 to March 2019. The website uptime for the Southern Asian sub-region averaged 98.2 percent during the time of the study while the average uptime for the 20 sites in this study was 93 percent. Although the Nepal NSO site (cbs.gov.np) studied by ODW was not included for this project, it may still serve as a proxy for how Nepal is doing in comparison to others in the subregion. The results below from the ODW study show that the NSO site in Nepal has a lower than the median uptime for NSO sites in the region.

Table 2: Regional uptime comparisons with other NSO sites in the region

Country	NSO URL	Time Online (%)
Sri Lanka	statistics.gov.lk	99.5
Bangladesh	bbs.gov.bd	99.4
Bhutan	nsb.gov.bt	99.2
Pakistan	pbs.gov.pk	99.0
Maldives	planning.gov.mv	98.8
Nepal	cbs.gov.np	98.7
Afghanistan	cso.gov.af	98.2
Iran	amar.org.ir	92.6

Note: ODW classifies countries in 20 regions, based on the United Nations Standard Country or Area Codes for Statistical Use (M49). The Southern Asia Sub-region differs from the World Bank's South Asia Region, in that it includes Iran, while the World Bank classification does not. Technical difficulties were experienced in the testing of India's uptime during this study and so their results are not included.

Map of site links

External links to a site are not only a source of traffic to a site, they are also important for a site to perform well in search engine results. The number of sites linking to the 20 sites in this study were evaluated. The results are displayed below in Figure 2. Eight sites had fewer than five domains linking to them and three sites only had one site linking to them: the Nepal Data Literacy Program page.

The sites in this study below are labeled and the dots surrounding them are websites that are linked to it. The sites with the most links were nepalindata.com; dataviz.worldbank.org; and nepalmap.org. The full list of the number of links to each site can be found in the final results spreadsheet.

Figure 2: Links to the 20 sites included in this study
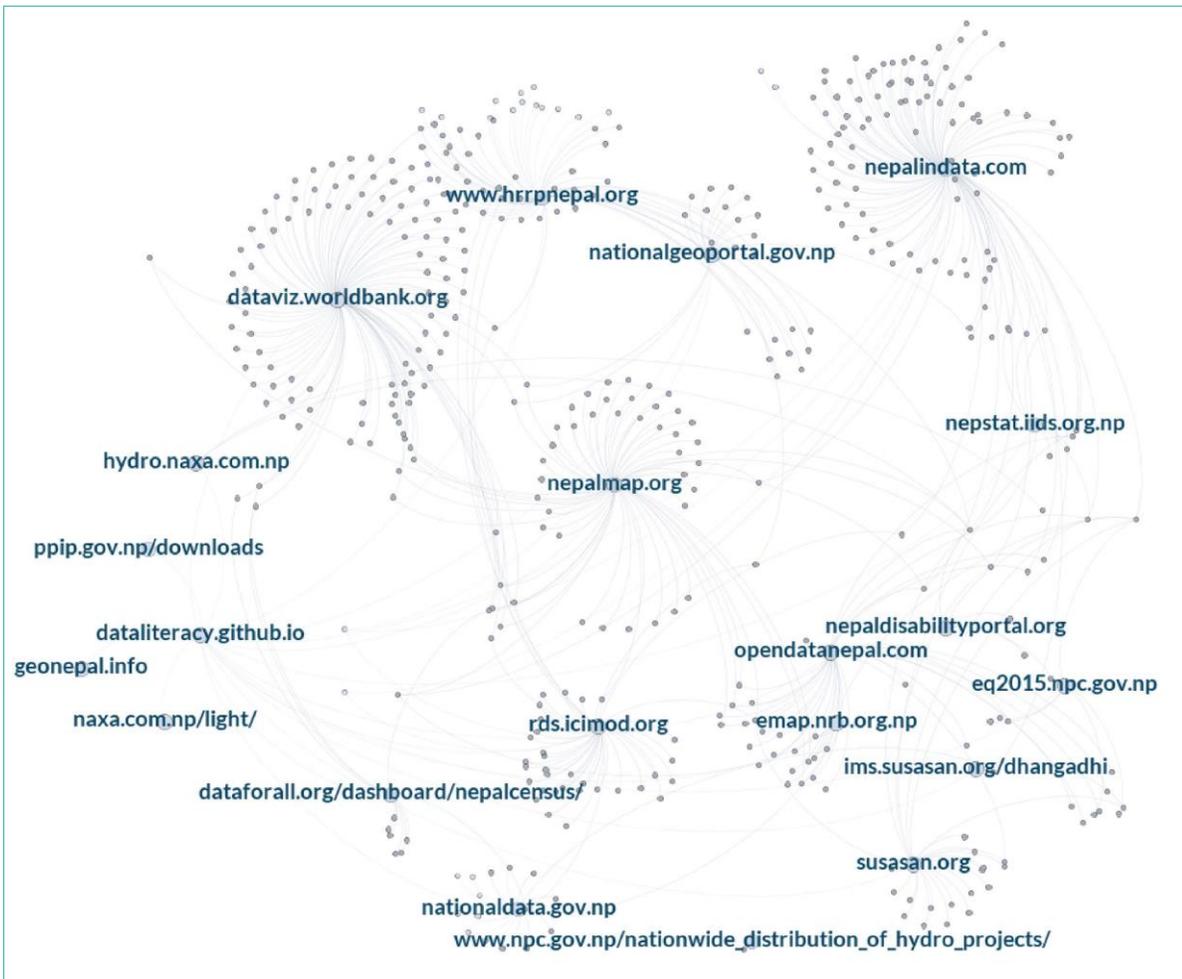
HOW ACCESSIBLE IS THIS SITE?

The sites tested had a much higher than average load time (2-3 seconds for the internet), which can be an impediment to use, especially in areas with weaker connectivity. However, all sites performed relatively well on the browsers tested and on mobile phones. Further, many sites were not available in multiple languages or were only available in English, which can limit their access to non-English speakers.

Aspects of the sites were also tested against web accessibility guidelines that are designed to make web content more accessible to people with disabilities. Eighteen of the sites presented issues for access by people with low vision or blindness. These findings were supplemented by an accessibility test through Google's Lighthouse program, a service that automates accessibility testing and evaluates dozens of factors related to accessibility. The following six sites received a failing score (below 70) from Google Lighthouse: EMapping System (40), Nepal Map (42), Nepal Human Development Indicators by District

(29), Nepal in Data: A Gateway to Development Data & Statistics in Nepal (62), Open Data Nepal (47), and CensusInfo Nepal 2011 (30).

CAN SOMEONE USE THE DATA ON THE SITE?

The majority of the sites tested had machine-readable file formats and acceptable download options, but many were missing metadata (data about data), so users may have a difficult time interpreting the data. Thirteen sites lacked a terms of use statement, which may cause uncertainty or deter someone from using the data.

Beyond these basic aspects of data usability, there is room to improve by providing more advanced download formats, such as JSON files (file format that facilitates the exchange of data). Additional options for bulk downloads on sites or application programming interfaces (or APIs, which allow for access to databases to update and access data) could also be added so that the data on the site can be used more quickly and for more advanced data functions.

Figure 3: Findings from the data usability assessment

Number of websites
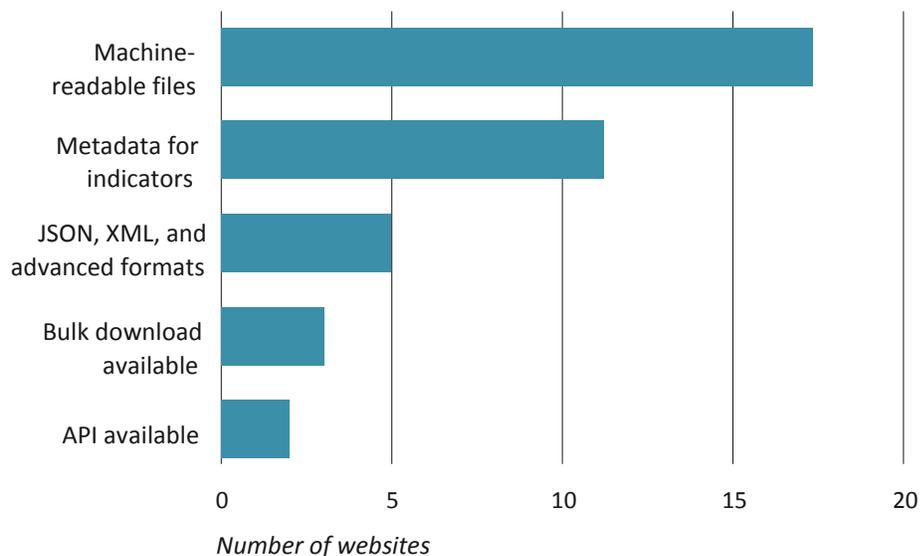
As interoperability is key to data usability, sites were also evaluated for their use of standard geographic nomenclature that reduces the difficulty of merging and joining datasets. While many sites evaluated contained subnational data, these data weren't identified by standard naming conventions for the level-one administrative units (province). According to the most recent Nepal Development Update (NDU) multiple public agencies in Nepal such as Central Bureau of Statistics

(CBS), the Survey Department, and the High-Level Commission for Information and Technology (HLCIT) have published administrative coding of government units.

Without standard names for administrative units, data users are forced to perform extra work to clean the data and get them ready to merge with other datasets, which can be a barrier to use. Given there are multiple naming conventions in Nepal, the NDU recommends a public agency such as CBS publish and widely disseminate a master data set of all geographic units and all naming conventions recognized by the government.

National Data Portal (nationaldata.gov.np)

To illustrate how the DSET evaluation results are performed at the data portal level, a summary of the results from the National Data Portal evaluation is provided below.

CAN SOMEONE FIND THE SITE?

The site is missing some key elements for SEO: XML site map, HTTPS URL, and connected social media accounts or blogs to enhance SEO and bring users to the site. The site does not have Google Analytics installed and experienced long periods of downtime during the evaluation.

HOW ACCESSIBLE IS THE SITE?

The site performed well on the accessibility test and had one of the highest scores on the Google Lighthouse accessibility metric (92). Improvements, however, could be made in the loading time (6 seconds) and the availability of user guides for the site.

CAN SOMEONE USE THE DATA ON THE SITE?

The site provides data in a machine-readable format, but improvements could be made in the availability of metadata and the provision of advanced file formats and bulk downloads option or provide APIs to access data.

RECOMMENDATIONS AND NEXT STEPS

This evaluation of sites in Nepal provides insight into the common accessibility, openness, usability, and functionality issues of data portals in Nepal. The findings can be used to provide training, technical assistance, or other support to address the most common issues. Recommendations are listed below.

Maintaining a data portal can be time consuming. More thought should be given to how these data portals can be made more sustainable. In ODW's research on website and data portal development, the team has found that sites not connected to larger domains perform worse on evaluations. These sites are often not connected to a larger ministry or organization and so struggle with resources and lack the site infrastructure provided by a larger domain. Larger domains generally have more traffic and data portals connected to them benefit as users from the larger sites are more likely to find them while on the site, thus increasing their traffic.

Recommendations for development partners, and portal sponsors and creators

Workshops, trainings, and technical support could be given to address the common weak points that were identified during this evaluation:

- Best practices for search engine optimization.

- Installing and leveraging Google Analytics to evaluate and increase data use.

- Reducing the downtime of websites and increasing their speed and performance.

- Implementing standard nomenclature for administrative units to increase the interoperability of data in Nepal.

- Improving and installing APIs and bulk download options on sites.

Serious thought should be given to the long-term management and sustainability of data portals that are not linked to larger sites or managed by ministries and larger organizations. Many of the issues found throughout this evaluation, and most notably uptime, could be addressed if these sites had more resources to address server issues and to implement best practices for data portals listed in the DSET evaluation.

Recommendations for data portal managers

Data portal creators and managers can take several steps to improve their sites' accessibility, openness, usability, and functionality. While the DSET evaluation found that each data portal had its own strengths, and areas for improvement, there are a few general recommendations that could improve many of the sites that were evaluated. These are provided below, and more details on which sites have implemented these features can be found in the full assessment spreadsheet online. For ease of use, these recommendations are grouped by the three main questions driving the DSET evaluation. Links to instructions on how to implement these recommendations are provided, where applicable.

To make it easier for a user to find their site, data portal managers can:

- Create an XML site map to make it easier for search engine bots to crawl their site.

- Install Google Analytics and use insights from the analytics to create a strategy for increasing website traffic and data use.

- Review the uptime of their site and make changes to the server configuration if average uptime drops below 99 percent.

- Publish blog content relevant to the search result keywords to increase ranking on search engines.

. Link their site to social media and post about updates to the platform, new datasets, and potential uses of the portal.

To make it easier for a user to access their site, data portal managers can:

- Improve the speed and load time of their site by: optimizing images, leveraging browser caching, serving scaled images, enabling compression, and minifying Javascript.

- Improve the accessibility of the site for people with disabilities by: increasing the contrast ratio between the background and foreground of the site, adding names and labels for attributes so that screen readers can read them, enabling zoom for all portions of the site, and making all the elements on the site accessible by using the tab key on the keyboard for users with motor disabilities.

- Create an option to view the site in both English and Nepali so that a wider variety of users can understand the content.

To make the data on the site more usable, data portal managers can:

- Publish metadata along with every dataset to provide context to data users.

- Prepare and publish a terms of use statement for the site to inform data users of their rights of use.

- Provide advanced options for bulk download of data on the site or provide APIs to access data.

- Furnish users with more advanced data download options, such as JSON formats.

Next steps to assess the effectiveness of data sites

To further support the creation of user-centric data sites and meet the data demands in the country, the following items should be reviewed:

1. The quality and reliability of data hosted on data sites.

2. The sources of the data hosted and whether the data are from primary or secondary sources.

3. The frequency at which data sites are updated.

4. The sustainability of the data sites and how the use of innovative technology and new revenue streams might improve their maintenance.

ANNEX

Data portals assessed

Table 3: Data portals studied in Nepal

#	Website link	Manager
1.	2015 Nepal Earthquake: Open Data Portal	Kathmandu Living Labs
2.	CensusInfo Nepal 2011	Central Bureau of Statistics
3.	EMapping System	Nepal Rastra Bank
4.	Housing Recovery and Reconstruction Platform (HRRP)	HRRP Nepal
5.	Hydro Map project	Niti Foundation
6.	Hydropower Projects in Nepal	Government of Nepal
7.	Infrastructure Management System Dhangadi	Dhangadhi Sub-Metropolitan City
8.	Kathmandu valley utility mapping initiative	Youth Innovation Lab and NAXA
9.	National Data Portal-Nepal	National Planning Commission, Central Bureau of Statistics
10.	National Geospatial Portal	Survey Department, Geographic Information Infrastructure Division (NGIID), Government of Nepal
11.	Nepal Disability Portal	D4D
12.	Nepal GeoNode	GeoNode
13.	Nepal Human Development Indicators by District	Kathmandu Living Labs
14.	Nepal in Data: A Gateway to Development Data & Statistics in Nepal	Bikas Udhyami
15.	NepalMap	Code for Nepal
16.	NepStat Database Portal	Institute for Integrated Development Studies
17.	Open Data Nepal	Open Knowledge Nepal
18.	Province Information	Government of Nepal
19.	Public Procurement Transparency Initiative in Nepal	Government of Nepal
20.	Regional Database System	ICIMOD
21.	SDG Sub-national indicators	World Bank Nepal
22.	Sustainable use of Technology for Public Sector Accountability in Nepal	Centre for International Studies and Cooperation (CECI)

Full list of data portal results

The full list of evaluation results for each of the sites above can be found on the GitHub site here.