

Executive Summary

This survey report intends to formulate an up-to-date picture of available and emerging open data resources on the National Committee for Sub-national Democratic Development (NCDD) budgeting process and opportunities to grow the open data community in Cambodia. The survey focused on identifying open data resources made available by the government, civil society, and universities. The survey process also undertook a mapping of existing open data platforms and data release cycles in Cambodia.

This survey was carried out based on the understanding that a healthy open data ecosystem is built on a supply (data producers) and demand (data users) equilibrium that rests on robust data infrastructures. On the demand side, the survey team evaluated open data availability and accessibility, how it was governed and the appreciation of open data among data producers. On the supply side, it considers how the data is used, what data gaps are perceived, and how the users appreciated and contributed to the open data community. For the data infrastructures, the team looked exclusively at functionalities of data portals/websites and how user-friendly they were.

Through the desk study more than 100 websites – those perceived as the most likely locations where users can find open data – were reviewed. For the supply and demand assessment, we interviewed 23 key informants, consisting of eight university lecturers or researchers, four civil society workers, four journalists, three freelancers and four government officers. These interviews were complemented by a publicly administered web-based survey that resulted in 79 responses. In total 102 individuals were surveyed, among which 26 were female.

Based on the analysis of the collected data, Cambodia's data spectrum in our sample tilts significantly towards consumption. Production accounts for only 20 percent of the respondents. Data dissemination was even less common, with only 13 percent of the respondents stating that their organization disseminates data. There is also a low understanding of the differences between data and information among data users and producers.

Open data at research institutions such as universities is virtually nonexistent even though researchers collect huge amounts of data. Four main reasons – intellectual property, no usefulness outside of the project, fear of violation of terms and conditions of the sponsors, and no trustworthy sharing platform – contribute to collected data remaining private. For civil society organizations, data production is almost always intended for project use or closed group sharing and is considered as having no public value or containing too much private information to be shared. However, often the results of the research are published in their project reports. At the national level, among more than 50 government websites reviewed, a large portion is information-based and does not provide the public with any data. When data is provided, it is almost always presented in an aggregated form such as tables, figures, charts or in-text numbers in PDF documents, which are sometimes just the scanned digital versions of the original printout. The government of Cambodia practically has no centralized data portal at all levels of governance. It is noteworthy that the UN has made more data about Cambodia available to the public than the Cambodian government itself.

In terms of usage, almost 90% of the respondents revealed that they used data from sources outside of their work to meet their data consumption needs, of which 82% is to support research, 74% report writing, and 47% project implementation. Although comprehensive assessment of open data literacy and attitude was not conducted, anecdotal evidence indicates

a low data literacy level and worrisome attitudes towards open data stewardship and usage. Only three percent of those taking part in this study correctly defined what constitutes open data. In parallel, all journalists interviewed intimated that they did not know what constitutes open data. They also claimed that their colleagues in the press sector were not trained to work with complicated data and lack functional data literacy. A large proportion of the questionnaire respondents (45%) do not know what metadata means.

The deficiency of enabling open data infrastructures and low open data literacy, combine to create an incredibly challenging ecosystem for the open data community. Information gaps, unreliable data due to poor data documentation and metadata, and low data quality are the top three perceived challenges faced by the data users. Real development challenges, however, come from both internal and external factors. As many countries in the ASEAN bloc are opening up their data repositories to their public and investing heavily on enabling data infrastructures, one would wonder how the Cambodian people will deal with a fast-changing, data driven digital economy and remain competitive. The dilemma surrounding these data driven changes for the Cambodian government centers around how well-prepared it is to invest in data production and delivery systems that ensure a reliable flow of data to the hands of its people who will then turn it into innovative and creative products and services.

The biggest data-induced problem that crosscuts every development aspect is that the limited availability of open data undercuts the public's oversight of the government's development initiatives. Government oversight through data is hard to achieve, except for the NCDD which provides the public some snapshots of the government's development initiatives through their nine public databases.

Despite these many issues, some opportunities for the growth of open data exist. An example is the World Bank funded Higher Education Institutions Capacity Improvement Project (HEICP), under which many scientific research projects will be supported. These projects will inevitably require a lot of open data and eventually will drive huge data growth. Another example is that the Cambodian government has begun to appreciate the importance of open data portals, as demonstrated by the launch of the Ministry of Environment's data portal. Some interests in voluntary geographic information systems were also observed.