Open Data Intermediaries for Non-State Actors in Cambodia

Feasibility Study

Dr. Pak Kimchoeun

Disclaimer

Table of Contents

List of Boxes, Figures, and Tables

List of Abbreviations

ADB	Asian Development Bank
ASEAN	Association of Southeast Asian Nations
CamDX	Cambodia Data Exchange
CDEI	Cambodia Development Institute
CSOs	Civil Society Organizations
CSV	Comma-Separated Values
EU	European Union
FGDs	Focus Group Discussions
IDB	Inter-American Development Bank
M & E	Monitoring and Evaluating
MEF	Ministry of Economy and Finance
MOI	Ministry of Interior
MOP	Ministry of Planning
MOU	Memorandum of Understanding
NGOs	Non-Government Organizations
NIS	National Institute of Statistics
NSDS	National Strategy for the Development of Statistics
NSS	National Statistics System
ODI	Open Data Institute
OECD	Organization for Economic Co-operation and Development
RGC	Royal Government of Cambodia
SWOT	Strengths, Weaknesses, Opportunities and Threats
The Foundation	The Asia Foundation
UK	United Kingdom
UNDESA	United Nations Department of Economic and Social Affairs
URI	Uniform Resource Identifier
WDR	World Development Report
WTO	World Trade Organization

Executive Summary

As Cambodia becomes more developed, policy interventions need to be based on reliable data to be effective. Promoting open government data is one part of the effort to improve data governance for policy purposes. Open data will not only help promote transparency within the government but also empower non-state actors to engage, provide feedbacks and hold the government accountable. In Cambodia, non-state actors have made only slow progress regarding data access and data use for policy purposes. One solution suggested by international examples for improving data sharing and use is to have data intermediaries.

This study seeks to assess the feasibility of setting up open data intermediaries for non-state actors in Cambodia and to provide key strategic considerations. The study builds on recent studies on public policy and the research sector in Cambodia. In addition, international literature was reviewed and interviews and focus group discussions (FGDs) conducted with international experts, key policymakers, NGOs, young researchers, and university students. A verification workshop was conducted on the preliminary findings of the study.

An open data intermediary is defined as an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of the open data that may otherwise not have been the case. It is important that prospective data intermediaries define and justify their roles especially in terms of how they can help address data sharing and use.

Experiences in other countries and in Cambodia suggest similar sets of challenges when it comes to data sharing and use. Those include lack of incentives to share data, commercial, ethical, risks associated with data sharing, and cost of data access and sharing. Risks associated with data sharing include legal and regulatory risks, ethical risks, reputational risks, and commercial risks. Several possible solutions can be considered to prevent and mitigate these risks, including anonymization, synthetic data, sharing under contract, engaging third-party data stewards, engaging the community, and describing and documenting the data.

Reflecting to the case of Cambodia, the study found that it is feasible to have open data intermediaries for non-state actors in Cambodia. This is because there are:

1. many opportunities in terms of a conducive policy environment,
2. areas for non-state actors to contribute,
3. willingness to collaborate from different stakeholders,
4. existing capacities (although with some limitations) and
5. the many low-hanging fruits that can be focused on in the short and medium-terms.

However, the feasibility also depends on how the prospective intermediaries manage to:

1. address their weaknesses, especially in terms of technical and human resources,
2. mitigate any risks associated with data sharing, and
3. gradually build trust among key stakeholders, especially with relevant government agencies.

There are two possible types of open data intermediaries to be considered. Based on the interviews, those are *1. Issue-focus open data intermediaries:* These refer to those organizations that focus on specific issues such as public budget transparency, local governance, social protection, education, health, agriculture, etc. The institutions best suited in these categories are the existing umbrella NGOs who work to assist their members and jointly advocate for policy changes in their areas of focus, and *2. Research-focus institutions:* This category refers to academic institutions, think tanks, and universities whose main focus is not advocacy in specific areas but promoting research and the knowledge sector in general. Their beneficiaries are not limited but include researchers, media, and the general public.

To build and operate effective and sustainable data intermediaries, long term commitment and strategy together with specific critical details are needed. Long term commitment and contributions are expected from the prospective open data intermediaries themselves, development partners whose financial and technical support are critical at least in the short and medium terms, and relevant government agencies whose collaboration is essential throughout the whole process. At the more operational level, the prospective data intermediaries need to pay more attention on getting the basic right from the starting phase, implementation phase, and sustaining and expanding phase.

1.
INTRODUCTION

1.1. Background and Rationale

As Cambodia becomes more developed, policy interventions need to be based on reliable data to be effective. This is particularly true as the Royal Government of Cambodia (RGC) is implementing its post-Covid-19 economic recovery plan and is preparing to successfully graduate upwards out of its lower income country status in 2027.[1] Thus, data needs to be adequate, reliable, and timely for policymakers to make necessary policy targets adjustments. This is very critical to ensure good use of public resources which have become scarcer due to external economic uncertainty.

Promoting open government data is often pitched by international transparency groups as a means to better use data for public policy purposes. International experiences point to specific benefits of open government data, including fostering of economic growth and job creation; improving efficiency, effectiveness, and coverage of public services; increasing transparency, accountability, and citizen participation; and facilitating better information sharing within government.[2] To realize these benefits, it is helpful to see open government data as an ecosystem comprising both the supply and demand side. The two sides reinforce one another.[3]

The data needed for policy purposes is generated through a national statistical system (NSS) consisting of various actors. In Cambodia (similar to other countries), the NSS includes the National Institute of Statistics (NIS), designated statistics units in various line ministries, and development partners who all are major data producers and users. Non-state actors are data users and therefore constitute a key element of the demand side in the open government data ecosystem. This group includes non-government organizations (NGOs), research institutions, academic institutions, researchers, university students, media, and the general public. These non-state actors are often seen as data users, but they too can be data producers.

In Cambodia, non-state actors have made only slow progress regarding data access and data use for policy purposes. International practices, as well as the RGC's policies, recognize the importance of non-state actors in accessing and using data to generate greater transparency and accountability

---

[1] (WTO, 2022) (RGC, 2021)
[2] (UNDESA, 2022)
[3] (IDB, 2018)

from public institutions. However, non-state actors in Cambodia have faced various challenges, including not only financial and human resources limitations, but also limited awareness and access to publicly available data, and limited capacity to analyze, interpret, and disseminate key findings. This is in contrast with the government, development partners, and the private sector, who, despite the many remaining limitations, have been increasingly better at producing more data, sharing, and using it among themselves.

One solution suggested by international examples for improving data sharing and use is to have data intermediaries. A data intermediary can help facilitate the collection, validation, and aggregation of data from data contributors and make data understandable, usable, and accessible to data users. For public data, data intermediaries can facilitate data sharing in a trusted, more efficient manner between government institutions or between government and non-government actors.[4] Data intermediaries also play valuable roles in the sharing and use of open data in developing countries.[5]

1.2. Objectives and Questions

This study seeks to assess the feasibility of setting up open data intermediaries for non-state actors in Cambodia and to provide key strategic considerations. It does so based on the context and rationale presented earlier. The findings and recommendations are intended for NGOs, research institutions, and academic institutions that might consider taking on the roles of intermediaries in their chosen fields or sectors, as well as for prospective government agencies and development partners whose collaboration and support are critical to make the data intermediation arrangements work.

In line with the research objectives, the study seeks to answer the four related questions presented below:

- From international literature, what are open data intermediaries? What are their roles and contributions to better data sharing and data use? What are the key strategic areas to pay attention to?
- Reflecting on the case of Cambodia, what are the rationales for having open data intermediaries for non-state actors? How feasible is the idea, including establishment of a think-tank association?
- If such data intermediaries are to be set up, what are the key strategic considerations that key stakeholders need to pay attention?

[4] (World Bank, 2021)
[5] (Francois, Canares, Chattapadhay, & Andrason, 2016)

- What are the specific tasks and issues that prospective data intermediaries need to consider undertaking?

1.3. Methodologies

The study builds on recent studies on public policy and the research sector in Cambodia. This includes the research commissioned by the Ponlok Chamnes program, such as the Diagnostic Study on the Policy Process and Use of Data- The Case of Cambodia and Insights from ASEAN (2022) and the Cambodian Youth and the Knowledge Sector – Challenges, Opportunities, and Recommendations (2022). As such, where relevant, the study will refer to these other studies, especially on relevant situational analysis, policy frameworks, and regional experiences.

In addition, additional international literature was reviewed and interviews with international experts on open data intermediaries were conducted. The concept on data intermediaries is relatively new in the Cambodian context, although what it entails is not entirely new from what some key government agencies and NGOs have already been doing. The nuanced difference is the emphasis on systematic and long-term strategic thinking that key stakeholders need to consider when setting up and supporting open data intermediaries.

Interviews and focus group discussions (FGDs) were conducted with key policymakers, NGOs, young researchers, and university students. The information was collected in addition to a review of the existing literature with a focus on their perspectives on the need for open data intermediaries for non-state actors in Cambodia, their expectation about the roles and benefits from the intermediaries, and their suggestions on how the intermediaries can be set up and supported, including their roles expanded and sustainability over time.

A verification workshop was conducted on the preliminary findings of the study. There were about 80 participants, including officials working for government agencies, NGOs, academic institutions, and young researchers. The discussions and feedback received at the workshop were used to improve the first draft, where relevant.[6] 2. Relevant concepts and experiences from other countries.

[6] The workshop was conducted on 24 November 2022 in Phnom Penh.

2.
RELEVANT CONCEPTS AND EXPERIENCES FROM OTHER COUNTRIES

A few key concepts and terminologies were used to guide the analysis and recommendations presented in this report. The first are the definitions of data and public data in particular. Based on the 2021 World Development Report (WDR), data is defined as 'information about people, things, and systems.' Data can be both quantitative and qualitative and can be stored on analog or digital media. Based on its intent, data can be either public or private purposes (i.e., for a commercial purpose). In this study, we focus on public intent data, also known as public data.[7]

For this report, the distinction among public data, government data, and open government data should be highlighted. Public data includes all data that are available in the public domain, including those created by Governments, academia, civil society, and the private sector. Government data – a subset of public data is "recorded and documented in any manner and any medium and …obtained and created upon the performance of public duties provided by law or legislation issued on the basis thereof." Lastly, Open Government Data is data open to and available in the public domain in various (including machine-readable) formats and normally licensed for all to access, use, modify, and share.[8]

While different organizations provide slightly different definitions, they are largely similar in terms of the key characteristics of open data. Comparing the definitions as provided by different organizations and high income countries,[9] open data should encompass the following five characteristics: free, publicly available, re-usabile, no restrictions, and machine-readabile. By 'free' and 'publicly availabile,' open data should be made available to the public. 'Re-usability' means all data should be reusable in whatever ways the users want. 'No restrictions' suggests that data should be freely available to everyone, while 'machine-readability' means data should be in a format that can be automatically read and processed by a computer. The degree to which a dataset is machine-readable is critical to its use. Because there is no detailed definition as to what makes data 'open' in Cambodia,

---

[7] (World Bank, 2021)
[8] (UNDESA, 2020)
[9] Examples of the different organizations and advanced countries that give definition on open data include Open Knowledge Foundation, Open Knowledge international, Open Data Barometer, Gartner, World Bank, OECD, the EU, USA, UK, Canada, Australia, and New Zealand.

this report refers to these characteristics when discussing open government data. Other conceptual frameworks and principles of open data are also useful as references (as summarized in the box below).

Box 1: Principles of open data and the 5-star deployment scheme

Government data shall be considered open if is made public in a way complies with the principles below.[10]

- **Complete:** All public data are made available. Public data are data that is not subject to valid privacy, security, or privilege limitations.
- **Primary:** Data are made available at the source, with the highest possible level of granularity, not in aggregate or modified forms.
- **Timely:** Data are made available as quickly as necessary to preserve the value of the data.
- **Accessible:** Data are available to the widest range of users for the wider range of purposes.
- **Machine Processable:** Data are reasonably structured to allow automated processing.
- **Non-discriminatory:** Data are available to anyone, with no requirement of registration.
- **Non-proprietary:** Data are available in a format over which no entity has exclusive control.
- **License-free:** Data are not subject to any copyright, patent, trademark, or trade secret regulation. Reasonable privacy, security, and privilege restrictions may be allowed.

Another helpful guide on open data sharing is the 5-star deployment scheme for Open Data developed by Tim Berners-Lee, the inventor of the Web and Linked Data initiator. Below is what each star rating means for open data.[11]

✶	Make your stuff available on the Web (whatever format) under an open license
✶✶	Make it available as structured data (e.g., Excel instead of image scan of a table)
✶✶✶	Make it available in a non-proprietary open format (e.g., CSV instead of Excel)
✶✶✶✶	Use Uniform Resource Identifier (URIs) to denote things, so that people can point at your stuff.
✶✶✶✶✶	Link your data to other data to provide context.

[10] (OECD, 2013)
[11] https://5stardata.info/en/

What is an open data intermediary? According to Schalkwyk et al (2016), an open data intermediary is defined as 'an agent (i) positioned at some point in a data supply chain that incorporates an open dataset, (ii) positioned between two agents in the supply chain, and (iii) facilitates the use of the open data that may otherwise not have been the case.' Schalkwyk et al (2016) also suggests that open data supply chains may comprise multiple intermediaries and that multiple forms of capital (namely, social, political, technical, and financial) may be required to connect the supply and use of open data. Here it is important to note that data intermediaries exist not just in the 'open public data' setting but also in the private sector world, where they play various roles ranging from data trust, data cooperatives, and data stewardship.[12]

Open data intermediaries need to define their roles and corresponding 'capital' to be effective. Because a country's open data ecosystem can be complex, it is important that prospective data intermediaries define and justify their roles, determine the 'capital' that they have and need to have, and the beneficiaries they intend to serve. These strategic level considerations determine the specific functions they can play and their potential partners.[13] Another way to determine such roles and contributions is to see how they can fit into the policy process, what challenges they seek to address in terms of data sharing and use, as well as what value-add they can bring to address those challenges.[14]

Several common challenges have been identified when it comes to data sharing and use, both in an open and non-open data setting. Those include:

1. lack of incentives to share data,
2. lack of knowledge,
3. commercial, ethical, and reputational risks,
4. legal and regulatory risks,
5. cost of data access and sharing, and
6. missed opportunities to use data in the public interest.

The Table below provides a brief explanation for each, together with the potential roles that data intermediaries can play to ease the challenges. It is important to note that which roles an intermediary should play, when, and how might vary largely on a case-by-case basis depending on the specific situation, opportunities, and constraints.

---

[12] (Janssen Heleen, Singh Jatinder, 2022)
[13] (Francois, Canares, Chattapadhay, & Andrason, 2016)
[14] (The Foundation, 2022)

Table 1: Issues that may prevent optimal data sharing and use[15]

Issues	Description	Data intermediary roles
Lack of incentives to share data	Data providers may not be sufficiently incentivized to share or provide access to their data (e.g., because sharing requires them to incur costs or effort that they are not able to recoup from those that benefit).	Operate a "pay to play" model where individual organizations can only access the larger datasets if they too have contributed data
Lack of knowledge	Data providers may lack sufficient knowledge of the potential uses of their data, while data users lack sufficient knowledge of what data could be made available and how.	Enable easier discovery of datasets, enabling data providers to generate value from their data and enabling data users to identify potential applications
Commercial, ethical and reputational risks	Perceived or actual risk of losing competitive advantage, suffering reputational damage from data uses that breach others' trust, or enabling ethically questionable uses of data may deter data access and sharing.	Provide assurance to data subjects and holders that data will only be accessed for defined purposes by approved individuals/organizations
Legal regulatory risks	Perceived or actual risks of breaching data protection, intellectual property rights, or regulatory requirements may also provide a deterrent to sharing.	Provide legal expertise and common data standards to facilitate legally compliant data sharing
Cost of data access and sharing	Costs may be prohibitive because of a lack of common foundations, infrastructure, and technologies that are needed for data sharing to be cost effective.	Provide data sharing environments that make effective use of technology including applying privacy-enhancing technologies
Missed opportunities to use data in the public interest	Cases where data sharing may be particularly likely to lead to economic and social benefits.	Provide secure access to sensitive data in a trusted research environment for the purpose of research in the public interest

[15] (CDEI, 2021)

Risks and lack of trust are the among the key hindrance to data sharing and use. Among the key challenges listed above, one area that tends to create mistrust, and thus discourage data sharing, is the limited attention to risks that are involved when sharing data. A recent report by the Open Data Institute (ODI) suggests that efforts need to be made to minimize risks when sharing data. Those risks include:

1. legal and regulatory risks,
2. ethical risks,
3. reputational risks, and
4. commercial risks.

A total of 13 guiding questions are suggested to assess these risks (see the Table 2). The paper also offers several possible solutions including:

1. anonymization,
2. synthetic data,
3. sharing under contract,
4. engaging third-party data stewards,
5. engaging the community, and
6. describing and documenting the data (e.g., the use of metadata documentation and catalogue).[16]

---

[16] (ODI, 2022)

Table 2: Assessing risk when sharing data[17]

Risks category	Description and assessment questions
Legal and regulatory	Perceived or actual risks of breaching data protection law, intellectual property rights, other regulatory requirements or legal contracts. • Does the data contain any personal data? • Does the data contain third-party data? • Do you have legal permission to share the data? • Are there any other relevant legal considerations?
Ethical	Perceived or actual risk of enabling unethical data collection or use, or of directly impacting people and communities. • Are there any relevant cultural considerations? • Is sharing the data likely to impact people or communities? • Will sharing the data impact the natural environment? • Does the data contain anything that could impact national security? • Does the data contain anything that could impact the security of the organization or its staff?
Reputational	Perceived or actual risk of suffering reputational damage from sharing or using data that breaches trust or that reveals limitations in processes or analyses. • Will anyone be surprised by you holding, sharing, or using this data? • Is a data-quality caveat required? • Are there any free-text or comment fields in the dataset?
Commercial	Perceived or actual risk of losing competitive advantage in the market. • Does the data contain anything commercially sensitive?

Digital platforms are key to effective data sharing in this digital age. With the advent of digital technology, the most effective way for data sharing is through the internet. As such, the level of digitalization, both on the government and non-government side, is considered a constituent part of the broader ecosystem of open government data.[18]

[17] (ODI, 2022)
[18] (UNDESA, 2022)

3.
AN OVERVIEW OF PUBLIC POLICY AND DATA USE IN CAMBODIA

The policy processes in Cambodia have been improved in the last five years but specific challenges remain. How the processes work – from formulation to implementation to monitoring and evaluation (M&E) – can be found in the recent report by the Foundation titled "A Diagnostic Study on the Policy Process and Use of Data".[19] Of note, in the last five years, standard policy processes have been established and continuously improved. An increasing amount of data has been used. However, it is still limited in terms of policy-budget linkages and M&E (especially at the impact level). Other de facto factors shaping the quality of the policy process include political interest, cross-ministerial coordination, and human factors (e.g., leadership, young technocrats, trust).[20]

Figure 1: Organizational framework for Cambodia's NSS[21]
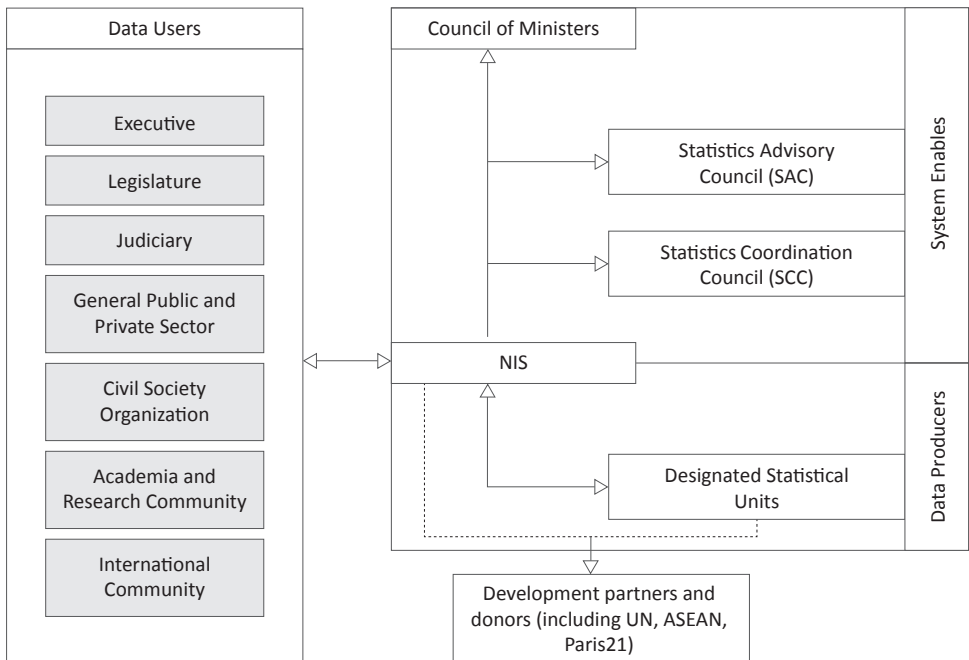
Effective data governance is recognized in Cambodia as critical to inform the policy process. As in other countries, the government of Cambodia and development partners are the biggest data producers and users. In the last decade, key stakeholders have produced large amounts of data in the NSS), namely, the NIS, line ministries, and other relevant agencies both at the national and sub-national levels. The government has also encouraged better use of data through various reform strategies, including the National Strategy for the Development of Statistics (NSDS) (2019-2023) and the Cambodia Digital Economy and Society Policy Framework (2021-2035). There has also been partnership at the ASEAN level to further improve the national statistical system and regional collaboration on data, including big data.

The NIS acts like the 'data intermediaries' for the whole NSS but it has been hampered by limited awareness and capacity constraints. In accordance with the Statistics Law (2015) and two sub-decrees defining the NSS in Cambodia, [22] the NIS plays a critical role in coordinating the data management of the government. However, in practice, there is limited awareness about the roles of the NIS and statistical units in line ministries and agencies. The government has tried to address these challenges by using digital technology and implementing the Data User Engagement Strategy, which is an important document, although still in draft form. The draft Data User Engagement Strategy sets out key objectives and principles which are in line with best open data practices, as summarized below.

[22] Those two sub-decrees are: 1) Sub Decree on Organization and Functioning of the National Statistical System (2007), and 2) Sub Decree on Designated Official Statistics (2010).

Table 3: Key points of the draft Data User Engagement Strategy[23]

Main objectives	Key principles	Means of engagement
• Create feedback mechanisms to enhance the understanding between users and producers in order to ensure the production of demand driven data and statistics • Enhance statistical literacy of users and policy literacy of data producers • Enhance coordination mechanism to improve accessibility, transparency, and availability of quality and reliable data in support of user engagement • Build mutual trust between producers and users of data to foster collaboration	• Inclusiveness • Accountability • Credibility • Confidentiality • Professionalism • Responsiveness • Harmonization • Standardization • Accessibility • Prioritization • Transparency/openness • Relevance	• Training, coaching, mentoring • Consultation workshops • User needs assessment • Inter-ministerial working groups • User-producer dialogues • Focus groups • High visibility events • Newsletters, press release • One-on-one engagement • E-communication platforms • Formal agreements

There has been only limited engagement from non-state actors due mainly to the persistent challenges of trust. Personal connection, friendship, and perceived political affiliations are still critical for trust-building between government and civil society organizations (CSOs). However, it is observed that it seems easier to build trust and connection among the younger generation of bureaucrats and young professionals in CSOs, partly because many of them tend to share educational and professional backgrounds (e.g., same schools, same scholarships and training programs, etc). Existing studies also identify a few other positive trends. First, in the last five years, Cambodian researchers have played more prominent roles in research institutions and think tanks. Second, there have been more young researcher groups formed and supported by various CSOs. The positive trends imply a strong potential for building the knowledge sector in Cambodia.[24]

The government of Cambodia has started to pay more attention to open government data, together with the idea of digital government. Although the notion of open government data is still new to Cambodia, this policy has been stated in various key government policies, the most noticeable of

---

[23] (NIS, Upcoming)
[24] (The Foundation, 2022)

which is the Digital Government Policy adopted in 2022. It has also created a National Digital Economic and Social Council, under which three committees are established:

1. The Digital Economy and Business Committee,
2. Digital Government Committee, and
3. Digital Security Committee.[25]

Per the Government's policies, various reform actions are planned to improve data governance, especially on the supply side. The Cambodia Digital Economy and Society Policy Framework 2021-2035 (2021) and the National Strategy for the Development of Statistics (2019-2023) point out key reform activities to promote data production, data quality assurance, data storage, and data sharing both within the government and also with non-state actors. [26] However, these reforms still need to move toward implementation.

---

[25] (RGC, 2022)
[26] (RGC, 2021) (RGC, 2018)

4.
RATIONALE AND FEASIBILITY ━━━━━━━━

4.2. Rationale

As mentioned earlier, for open government data to be effective, improvements in both supply and demand sides need to move in tandem. While there are noted reform activities on the supply side, this report zooms in on the demand side, especially that related to non-state actors. The engagement of non-state actors in the whole data ecosystem has been limited, lagging behind key government actors and development partners. This is the rationale leading to the question of whether data intermediaries are needed to assist non-state actors.

All the key informants and participants in the study strongly support the idea of having data intermediaries for non-state actors. Again, those non-state actors include NGOs, academic institutions and universities, researchers (especially young researchers), the media, and members of the public. The respondents noted that they expect the data intermediaries can help address and overcome specific needs and gaps on both the non-state actors and Government side.

On the non-state actor side, a few challenges have been persistent. Those include:

1. limited access by NGOs, academic institutions and universities, researchers, media and the public to the public data needed to improve their works either for advocacy or other purposes,
2. limited awareness among these non-actors about the already available data,
3. and their limited capacity on data interpretation and visualization.

On the government side, there has been an explicit willingness to improve data sharing and use. This was noted in the interviews conducted and is also included in the written policies, including the National Strategy for the Development of Statistics (2019-2023) and the User Engagement Strategy.[27] As mentioned earlier, the User Engagement Strategy, although still in final draft form, includes an objective that in line with the ideas that data intermediaries try to achieve. That objective is to:

---

[27] (RGC, 2018) (NIS, Upcoming)

- Create feedback mechanisms to enhance the understanding between users and producers to ensure the production of demand-driven data and statistics,
- Enhance statistical literacy of users and policy literacy of data producers,
- Enhance coordination mechanism to improve accessibility, transparency and availability of quality and reliable data in support of user engagement,
- Build mutual trust between producers and users of data to foster collaboration.

As data users, non-state actors have lagged in their data access, awareness, and capacity to use data. Based on the previous studies and the key informant interviews conducted for this study, it is confirmed that the government and development partners have made much more noticeable progress in data sharing among themselves, leaving other non-state actors out of this data sharing.

Specific examples of the progress made by the government are the use of the CamDX system and CamStat in the government. But improvements are still needed in terms of data quality, accessibility, user-friendliness, and regularity of the data sharing.[28] Even with these limitations, these initiatives represent a huge advance compared to the non-state actor side. For instance, with the CamDX, the exchange of data across ministries involved in, say, business registration, can be done through a digital platform which is much more efficient and less prone to human errors. Relevant government agencies (especially the NIS, MEF, and others) have recognized the limitations and work to improve them through various ways, one of which is to better engage and get feedback from non-state data users.

The limitations on the part of the non-state actors have led to many missed opportunities for them to contribute to the policy process. From the previous studies and the interviews, both the NGOs and the government agree that there has been growing space for non-state actors to engage and provide input into the policy process. However, due to the limited data access and related capacity constraints, the space has not been well used.

The limitations also hinder the long-term development of human resource capacity in data-related issues. Such capacity is needed not only by the non-state actors but also the government. Interviews and policy documents indicate that to move to digital economy, digital government, and digital society, more people with data and digital skills are much needed. Currently,

[28] (The Foundation, 2022)

such resource exists only in scarcity. Yet, existing studies and interviews suggest that there has been more interest from academic institutions, universities, and young researchers to promote and engage in research works. However, limited access, awareness, and capacity relating to data have acted as persistent hurdles for them to realize their potential.

In short, there is a real need for open data intermediaries for both short and long-term development priorities in Cambodia. The open data intermediaries can play roles to not only address the current challenges and gaps but also to contribute to the longer-term priorities of the government to improve the overall national statistical system. They can also help to build the digital Government, economy, and society in Cambodia in the medium and long-term.

4.3. Feasibility

It is feasible to have open data intermediaries for non-state actors in Cambodia. Based on the literature and the opinions of key informants, this paper inserted the information into a SWOT (strength, weakness, opportunity, and threat) framework as follows:

- It is feasible to set up and operate open data intermediaries because there are:

  1. many opportunities in terms of a conducive policy environment,
  2. areas for non-state actors to contribute,
  3. willingness to collaborate from different stakeholders,
  4. existing capacities (although with some limitations) and
  5. the many low-hanging fruits that can be focused on in the short and medium-terms.

- However, the feasibility also depends on how the prospective intermediaries manage to:

  1. address their weaknesses, especially in terms of technical and human resources,
  2. mitigate any risks associated with data sharing, and
  3. gradually build trust among key stakeholders, especially with relevant Government agencies.

Figure 2: SWOT for the feasibility of open data intermediaries

Strength	Weaknesses
• Existing engagement between non-state actors and state actors • Existing capacity among non-state actors • Existing 'trust' between some non-state actors, government agencies, and others • Existing work to promote data sharing	• Limited technical and financial capacity to perform necessary roles as open data intermediaries • Still limited trust between non-state actors and Government • Limited long-term institutional commitment to perform and build capacity
Opportunities	**Threats/risks**
• Conducive and encouraging policy environment • Specific areas for contribution by non-state actors, especially on M&E and feedback • Willingness to collaborate from different stakeholders • Existing progress and large amount of data already made publicly available • Potential young researchers with capacity to absorb new technical knowledge, including digital technology	• Limited attention to the question of risks when sharing data and how to mitigate them • Limited appreciation from development partners to support data related capacity (especially for long-term)

Long-term commitment and investment are critical to the feasibility of effective and sustainable data intermediaries. Given the need for and expected roles of open data intermediaries and the SWOT as presented above, a few key points should be considered at the design and implementation stages so that the to-be-data intermediaries can not only perform but also sustain their roles and make contributions over the long-term. This report includes suggestions in the next section. To be emphasized here is that the nature of the work performed by open data intermediaries requires long-term commitment and investment.

The idea of setting up a Think Tank Association is also feasible but should be seen as a signpost to reach in the medium term. In the discussion that led to this study, a question was raised about whether it is necessary and feasible to set up and operate a Think Tank Association. On this question, our key informants appreciate the idea but suggest that such a forum should be pursued in the medium term as discussed in the next section.

5.
RECOMMENDATIONS FOR NEXT STEPS ━━━━━

To build and operate effective and sustainable data intermediaries, an overall strategy with specific critical details is needed. This recommendation section covers ideas for the overall strategy, which various stakeholders can consider if they are to support the setting and functioning of open data intermediaries.

5.1. Overall Strategic Considerations

Long term commitment and contributions from all key stakeholders are needed to make open data intermediary arrangement work in Cambodia. Those include primarily:

1. the prospective open data intermediaries themselves,
2. development partners whose financial and technical support are needed at least in the short and medium terms, and
3. relevant government agencies whose collaboration is essential throughout the whole process.

There are two possible types of open data intermediaries to be considered. Based on the interviews, those are:

- **Issue-focus open data intermediaries:** These refer to those organizations that focus on specific issues such as public budget transparency, local governance, social protection, education, health, agriculture, etc. The institutions best suited in these categories are the existing umbrella NGOs who work to assist their members and jointly advocate for policy changes in their areas of focus, and
- **Research-focus institutions:** This category refers to academic institutions, think tanks, and universities whose main focus is not advocacy in specific areas but promoting research and the knowledge sector in general. Their beneficiaries are not limited but include researchers, media, and the general public.

There are two types of government agencies expected to be involved. Based on the interviews, those are:

- **Central ministries/agencies** with a direct mandate in public data governance such as the National Institute of Statistic /Ministry of Planning (NIS/MOP), MEF, Ministry of Interior (MOI), and Ministry of Post and Telecommunication (MPTC), and

- **Line ministries/agencies** who produce and share data on specific sectors (e.g., education, health, agriculture, investment, trade).

This overall strategy is informed by the SWOT analysis presented above. The strategy should:

- Maximize harnessing and expansion of opportunities
- Minimize negative impacts and mitigate current and potential risks
- Focus on leveraging and improving upon strengths, and
- Focus on building internal capacity to address existing gaps and areas of weaknesses.

Based on the proposed strategies, there are several areas that each of the three identified groups of actors can focus on. Those are summarized in the Table below.

Table 4: Strategic areas for each actor to focus on

Key actors	Specific areas/tasks to focus on
Prospective open data intermediaries (See more details in the next section)	• Identify policy areas that they should focus on to justify their roles and contributions • Develop long-term plans to develop and sustain themselves as open data intermediaries • Establish relationships (preferably formal) with relevant state agencies and other stakeholders to build trust and mitigate risks • Offer specific support to potential government partner agencies to share some workloads needed to make data sharing more possible • Identify potential development partners and reach out for potential support
Government agencies	• Open up possibilities of having formal collaboration with interested non-state actors acting as open data intermediaries • Provide the necessary support to the open data intermediaries, including information on relevant policy needs • Make data sharing more open and accessible in accordance with the existing legal framework and international practices • Share concerns and find joint solutions with the open data intermediaries on how to best mitigate risks when sharing data • Allow for continuous dialogues for learning-by-doing with the open data intermediaries
Development partners	• Prioritize funding for data projects, especially those to be conducted by non-state actor open data intermediaries • Act as a 'bridge' to help build trust and encourage more data sharing between state actors and non-state actors • Link actors working on data in Cambodia (both state and non-state) with regional and international networks for broader knowledge sharing • Make data collected more easily available

5.2. Specific suggestions for open data intermediaries

This report suggests several recommendations for prospective open data intermediaries on how to set up, operate, sustain, and expand their activities in the near and long terms. The following suggestions are intended to be issues- or tasks-specific and practical, which can be included in a more coherent strategic framework/results framework.

Table 5: Specific suggestions for prospective open data intermediaries

Specific issues/tasks	Suggestions
1. During starting phase	
Identify long-term objectives	• The objective of the open data intermediaries should reflect their intended roles and be in line with the government's policy (e.g., the User Engagement Strategy)
Identify the areas of engagement	• The areas of engagement should be in line with their existing areas of focus and their strengths • The selected areas of engagement should also be justified by showing the potential for better data use

Identify the policy questions (and indicators) to focus on	• Within the broader areas of engagement, specific policy questions need to be zoomed in to allow for more targeted data needs
	• The relevant policy questions can be identified from:
	  1. sectoral policy documents including the Budget Strategy Plan (BSP),
	  2. relevant SDG indicators and sub-indicators, and
	  3. most importantly, the M&E framework for each sector.

Identify data needs and map available data	• Specific needed data sets and data points should be identified to answer the selected policy and information needs • A mapping exercise should be done to identify what data already exists, from where, in what formats, and to identify any remaining gaps • Review M&E frameworks of rthe elevant sectors to answer these questions
Seek formal collaboration with relevant government agencies	• Relevant government agencies can include line ministries/agencies in charge of specific sectors and/or central ministries/agencies who work directly on data issues. Ideally, data intermediaries should collaborate with both types of government agencies to ensure they have a good grasp of both sector issues and data governance issues. • Formal collaboration should be developed at the very beginning of the process to ensure formal buy-in, trust, and credibility with potential funders • The relationship should be formalized either based on an existing or a new MoU. The MoU should address specific areas to focus on, data to use, and key principles to mitigate any risks associated with data sharing
Identify technical and financial capacity needs and potential funding support	• Needed technical and financial support should be identified and potential funding proposals prepared • Potential funding support can come from donor agencies already funding the sectors or other regional and international donors

2. During the implementation phase	
Make use of the already publicly available data	• Before asking for more data to be shared, it is important that available data is analyzed and used • Data intermediaries should work and consult with their partners in the government to identify how they can make good of those existing resources and digital platforms (including the CamStat)
Advocate for public sharing of the more user-friendly data format	• While maximizing the use of available data in PDF, data intermediaries can in parallel advocate for sharing of those same datasets in Excel format
Use digital platforms to share the data, designed with user-friendliness in mind	• Effective data sharing needs to be done online and in a user-friendly format which requires investment to develop and maintain
Invest in necessary effort to turn existing key datasets from PDF to Ms Excel format	• While advocating with the government to release Ms Excel data, data intermediaries should be willing to invest some resources (e.g., hiring more staff and interns, developing databases or websites) to convert key datasets currently available in PDF to Ms Excel format and make them available on the web

Establish dashboards to 'tell stories' from the data	• Having dashboards on the websites is a good way
	  1. to demonstrate why data sharing is useful and
	  2. to show simple stories emerging from the collected data. This can generate 'positive demonstration effects' for both data sharers and data users

Consider specific terms and conditions of data use from the website	• While making data available online, specific terms and conditions should be attached to protect the posting organizations from any unforeseen risks • Data intermediaries should refer to any relevant legal provisions in Cambodia (to be identified) or practices by international organizations (such as the World Bank, the Asian Development Bank (ADB)) to develop such terms and conditions[29]
Engage youth in the process (especially on areas relating to digital technology)	• Data intermediaries should recruit youth either as full time, part-time, interns, or volunteers to help with the data works, ranging from data entry, online surveys, data sharing, website development, etc.
Provide regular updates and training on data availability awareness, data analysis, and data visualization	• Data intermediaries should see it as their role to provide regular updates and training on data availability, data analysis, and data visualization, as possible within their resources.
Solicit regular feedback from data users	• Develop specific mechanisms to receive regular feedback from data users, be them their members (for membership NGO) or researchers in general
Interact regularly with the Government agencies to provide inputs and feedbacks	• Facilitate regular interactions with relevant government agencies which are needed to deepen trust and mutual understanding, especially over questions of risks and the usefulness that comes from the data sharing.

[29] For an example of the ADB's terms of use for the datasets shared on its website, please visit: https://data.adb.org/terms-use-data

3. Sustaining and expanding	
Demonstrate effectiveness through learning by doing	• A practical but effective approach of learning by doing is strongly recommended. The approach can be manifested through regular outputs, feedback, and regular adjustments among key stakeholders involved • The feasibility of sustaining and expanding the works of open data intermediaries rests heavily on their ability to create 'demonstration effects' to convince relevant government agencies and development partners to see the values of their roles and contributions.
Consider collaboration with other non-state actors (including the creation of the Think Tank Association)	• As trust gets built and individual data intermediaries have proven their trustworthiness, collaboration can be considered with specific areas of contribution and accountability structures. • Should it be considered helpful, an association can be established among several research institutions and data intermediaries.
Consider taking on more advanced functions as data intermediaries	• As their capacities advance, data intermediaries can go beyond their focus on 'open government data' to 'data' in general. In such case, they can consider taking on more functions as data intermediaries such as data trusts, data exchanges, personal information management systems, industrial data platforms, data custodians, data cooperatives, and trusted third parties.[30]

[30] (CDEI, 2021)

References

- CDEI. (2021). *Unlocking the value of data: Exploring the role of data intermediaries. London.*
- Francois, S., Canares, M., Chattapadhay, S., & Andrason, A. (2016). *Open data intermediaries in developing countries. The Journal of Community Informatics, 1721-4441.*
- IDB. (2018). *Toward an open data demand assessment and segmentation methodology. Washington DC.*
- Janssen Heleen, Singh Jatinder. (2022). *Data intermediary. Internet Policy Review.*
- NIS. (2021). *Draft Data User Engagement Strategy. Phnom Penh.*
- NIS. (Upcoming). *User Engagement Strategy (Draft). Phnom Penh.*
- ODI. (2022). *Assessing risk when sharing data: A guide. London.*
- OECD. (2013). *Open government data: Towards empirical analysis of open government data initiatives. OECD.*
- RGC. (2018). *National Strategy for the Development of Statistics (2019-2023).*
- RGC. (2021). *Cambodia Digital Economy and Society Policy Framework 2021-2035. Phnom Penh.*
- RGC. (2021). *Strategic Framework and Program for Cambodia's Economic Recovery (2021-2023).*
- RGC. (2022). *Digital Government Policy. Phnom Penh.*
- The Foundation. (2022). *A Diagnostic Study on the Policy Process and Use of Data: The Case of Cambodia and Insights from ASEAN, Phnom Penh. Phnom Penh.*
- UNDESA. (2020). *E-Government survye for 2020. New York.*
- UNDESA. (2022). *E-Government survey 2022: The future of digital government. New York.*
- World Bank. (2021). *World Development Report 2021: Data for Better Lives. Washington, DC.*
- WTO. (2022). *Trade impacts of LDC graduation - Cambodia.*

Ponlok Chomnes is a four-year initiative (2019-2023) to strengthen the capacity of the knowledge sector and inform public policy analysis and dialogue in Cambodia. In partnership with the Australian Department of Foreign Affairs and Trade, The Asia Foundation is strengthening organizational and technical capacity among Cambodian research institutions and creating an enabling environment for policy dialogue.

CONTACT US