Global Partnership
for Sustainable Development Data

Unlocking Privately Held Data for Public Good
Real time collaborative learning for COVID-19 and Beyond

Background

The Covid-19 pandemic has raised the stakes around public-private data sharing. Governments are rapidly establishing arrangements with mobile and other companies to access mobility and location data to support contact tracing and to assess the impact of various mitigation measures[1,2,3]. At the same time, digital rights activists are sounding the alarm bell, seeing a range of disparate efforts where they worry that consideration for individual rights is uneven and insufficient attention has been given to the risks of entrenching a surveillance state.[4,5,6]

Many public-private data sharing experts are lamenting that we have not learned the lessons of the past. After years of pilots, research, workshops and conferences, the major political, economic, ethical, legal and technological obstacles to unlocking privately held data for public good are well known. But this has not translated into an established and reliable system or set of mechanisms and tools for facilitating rapid, frictionless and trusted public-private data sharing for the Covid-19 response at scale. There are several important efforts underway to address this gap that have been in development since before the pandemic, but they have not yet reached their full operational potential[7,8].

Even so, the COVID-19 pandemic is demanding rapid action and many experts in this field are working with governments on crisis response initiatives that are using privately held data for real time decision making. The unique circumstances of this moment offer an opportunity to learn alongside these initiatives, build on past experience and knowledge, examine how the context for public-private data sharing is changing, and work towards scalable solutions for the future.

This initiative builds on our earlier thinking supported by the Hewlett Foundation about how to foster a collaborative learning agenda to better understand the various models for unlocking

---

[1] The Economist, Countries are using apps and data networks to keep tabs on the pandemic, https://www.economist.com/briefing/2020/03/26/countries-are-using-apps-and-data-networks-to-keep-tabs-on-the-pandemic

[2] Caroline Buckee et.al, Aggregated mobility data could help fight COVID-19, https://science.sciencemag.org/content/368/6487/145.2

[3] Joseph Bullock et.al, Mapping the Landscape of Artificial Intelligence Applications Against COVID-19, https://arxiv.org/pdf/2003.11336.pdf

[4] Sean McDonald, The Digital Response to the Covid-19 Outbreak, https://www.cigionline.org/articles/digital-response-outbreak-covid-19

[5] The Hill, How human-centered tech can beat COVID-19 through contact tracing, https://thehill.com/opinion/technology/493648-how-human-centered-technology-can-beat-covid-19-through-contact-tracing

[6] Politico, In fight against coronavirus, governments embrace surveillance, https://www.politico.eu/article/coroanvirus-covid19-surveillance-data/

[7] Nuria Oliver et.al, Mobile phone data and COVID-19: Missing an opportunity? https://arxiv.org/pdf/2003.12347.pdf

[8] Stefaan Verhulst et.al, Mapping how data can help address COVID-19, https://medium.com/data-policy/mapping-how-data-can-help-address-covid19-a7be2e631aec

privately held data for public use, and the implications for taking them to scale. Our aim was and still is to surface lessons and identify ways to overcome persistent barriers by facilitating frank, in-depth learning among the leaders in this field being guided by the needs and challenges they see to moving the field toward scalable, sustainable and responsible data use.

Objective

We will create a space for collaborative learning across leaders in the public-private data sharing space. Our aim is:

- To create immediate connections and opportunities for sharing learning across COVID-19 response efforts, offering key players in this space an efficient and light-touch way of learning what others are doing and hashing out common challenges.
- To explore what we can draw from the COVID-19 experience to address the commonly known barriers in this field, examine how the context may be changing, and pave the way for smoother, scalable, safe and trusted private data use for future emergencies and recurrent data use in stable situations.

Our overall goal is to work towards the safe and trusted use of privately held data by public actors by engaging with the features of the predominant existing operational models – distinguished principally by different data access mechanisms[9]**.** Focusing on the operational models provides an entry point for examining several key questions about the characteristics and implications of individual initiatives as well as about the greater community of practice and context in which models are developed. There has been less analysis comparing the functioning and impact of different operational models, how they contend with the legal, ethical, economic, and capacity challenges inherent in public-private data sharing, and what this means for their suitability to different environments. This focus may evolve along the way.

Approach

We will ensure that the learning process is underpinned by thoughtful preparation, that virtual engagements are structured and carefully facilitated, and that the material from each engagement is digested and reflected back to participants in a timely manner and in a way that informs each subsequent engagement. The preparation and facilitation will be designed to respond to the needs and priorities of participants to maximize its usefulness to them and to the broader community. We will strive to use a variety of innovative platforms for engagement and, depending on the topics and mix of participants, this may mean splitting discussions into smaller

---

[9] *Sample typology of access mechanisms from 2020 Final report prepared by the High-Level Expert Group on Business-to-Government Data Sharing: Towards a European strategy on business-to-government data sharing for the public interest – Technically speaking, data could be accessed as follows: Remote access, Q&A mechanism, Limited release, Pre-computed indicators and synthetic data, https://ec.europa.eu/digital-single-market/en/news/experts-say-privately-held-data-available-european-union-should-be-used-better-and-more*

groups. Above all else, we will aim to keep the burden on participants low and strive to be sensitive to the demands of the current crisis situation.

The first phase of this learning process will be structured around four touchpoints. The first will aim to rapidly scope key questions and topics of interest to the participants and the following touchpoints will dig into these topics. With each touchpoint we will assess the usefulness and viability of the approach and rely on participants' advice on how to make it as meaningful as possible. As we move from one touchpoint to the next, we will assess whether and how to build out a second phase to facilitate additional connections, deepen the learning and/or spark concrete collaborative work on specific problems.

Touchpoint 1 - Scoping

The scoping will aim to foster a collaborative spirit from the start and ensure that the learning process is responsive to the different perspectives, experiences and interests of participants. This will include a preparatory survey of participants, key informant interviews with a select number of participants and a virtual discussion lasting 2-3hrs. The main goals of this first touchpoint will be to build a common vision for what the learning process will achieve and the key elements it should cover.

Touchpoints 2-4 – Digging in

We will structure the subsequent touchpoints around specific questions or issues that participants prioritize for deeper analysis and sharing (see below for a preliminary list). Through these engagements we will create opportunities for participants to share and showcase what they're learning from Covid-19 response projects in real time as well as bring in participants' past experience and expertise to broaden the discussions. As already noted, we will draw on different platforms for virtual engagement and will likely divide into smaller groups for more intimate discussions and then share with the broader group.

We will document and share the learnings from the first phase and produce other practical tools that may emerge from the discussions and collaborative work.

Timeline

Apr 27 – May 15	Participant survey and key informant interviews
May 21, 8:30am-11:30am EST	Touchpoint 1 – Scoping
Week of June 1 (time/date TBD)	Touchpoint 2
Week of June 15 (time/date TBD)	Touchpoint 3
Week of July 6 (time/date TBD)	Touchpoint 4

Participants

Participants will include a mix of key stakeholders involved in developing, implementing and analyzing operational models for unlocking privately held data for public good, including:

- Third party organizations that have developed and implemented operational models for public-private data sharing working with both data holders and users
- Government and other public agency data users
- Private data holders that have shared data for public sector use
- Analysts and thought leaders on public-private data sharing

We will not restrict participants to those involved in Covid-19 efforts because we want to bring in as much relevant experience for both the real-time learning around crisis response and the longer-term questions on recurrent data use. The virtual formats will allow us to be flexible and perhaps engage additional participants for certain topics.

The learning process will aim to provide value to participants by
- Maximizing their role in designing a process that meets their needs;
- Providing the opportunity to connect on related COVID-19 efforts and share learning that may be immediately applicable to those efforts;
- Drawing out lessons from the COVID-19 experience to work collaboratively on concrete problems that will facilitate smoother data sharing partnerships in the future.

Potential learning questions

This is a preliminary list that will be shaped by participants

- If the goal is safe use of big data by government actors, to what extent does this depend on features inherent to the data sharing models themselves versus external factors, or the interplay between the two?

- How do different models facilitate public trust through oversight, decision-making and capacity building structures? Are different measures needed for non-emergency situations?

- What are the accountability mechanisms for decisions around privacy preserving techniques? Who creates and governs oversight or ethical frameworks?

- To what extent do the different operational models lend themselves to incentives that encourage sustained engagement from all partners, private and public?

- Is it possible to have standardized metrics to evaluate how, when, and how well models work? To compare value-add of different models in specific contexts?

- How well does each model adapt to different capacity and resource levels among data holders and data users? How well do they adapt to different policy and regulatory environments?