



**Global  
Partnership**  
for Sustainable  
Development Data

# Unlocking Privately Held Data for Public Good

Real time collaborative learning  
for COVID-19 and Beyond



---

**Final Report**  
July 2020

# Public-private data sharing in the time of COVID-19: Why a collaborative learning series?

COVID-19 response efforts that use privately held data to inform government decisions have come together rapidly. The urgency of the crisis enabled these initiatives to circumvent many of the well-known barriers that historically held back the use of private data for public good at scale. But, urgency has also led to partnerships being established without firm policy, regulatory, capacity and technical foundations.

Rather than waiting for the crisis to end to learn from this experience, we saw an opportunity for practitioners to learn alongside ongoing initiatives. With support from the William and Flora Hewlett Foundation, we created a four-part virtual series to share learning from current response efforts and to identify lessons that might help address the commonly known barriers to public-private data sharing, and pave the way for more scalable, sustainable and responsible private data use in the future.

## MAIN TAKE-AWAYS:

# Persistent barriers for users

### **The combined learning of participants in the series points to significant barriers on the user side of data sharing initiatives.**

Partnerships have been established quickly, and users in the public sector are struggling to find guidance and advice that meets their needs across a range of technical and governance challenges. And yet, there is a vast landscape of existing tools, suggesting that there is a need to better understand which tools and guidance are actually used and why.

Companies and governments are finding that they still cannot agree on financially sustainable models to sustain partnerships because the demand is not strong enough. Decision-makers in government are increasingly seeing the value of privately held data, but until insights are delivered in terms of what problem it will help solve or what service it will help deliver, busy decision-makers will not be willing to invest.

There is wide support for efforts to build a common infrastructure and standards that can facilitate data re-use at scale, but these are long-term aspirations that risk being disconnected from immediate user needs.

## CONTEXT:

# The state of public private data sharing

Meeting the Sustainable Development Goals requires bringing different data sources together for new, more granular and timely insights, and strengthening the capacity of decision-makers to use data more effectively. A critical element of this involves unlocking the data held by private firms and sharing that data for the public good. More than a decade of experience with different technical, governance and partnership models for public private data sharing has shown the benefits for public policy and planning. But, persistent barriers are holding back public private data sharing at scale.

Initiatives to build data infrastructures,<sup>1</sup> platforms and tools,<sup>2</sup> and networks<sup>3</sup> to facilitate more scalable, sustainable public-private data sharing models have begun to make significant progress in the face of these challenges. But the community of practice continues to work in a fragmented manner, and sustainability with respect to economic and political considerations remains a challenge.

## BARRIERS TO PUBLIC PRIVATE DATA SHARING

- Lack of common standards limit interoperability and quality of insights
- Weak legal and regulatory frameworks contribute to a lack of trust
- Low capacity on the data holder and data user sides limits the scope for analysis and an understanding of what is possible and limits transparent communication with the public around management of privacy and security
- Misaligned incentives and a limited body of evidence on the value of data prevents partners from agreeing on business models based on shared value

## CONTEXT:

# The case of COVID-19

COVID-19 is exceptional in the speed and breadth of its spread, necessitating response measures that have raised concerns around data protection and privacy. Governments around the world are rapidly rolling out tools and technologies that are testing the limits of privacy and data protection regulations, or being launched in environments where such rules are weak or insufficient.

The challenge is often framed as a dichotomy between privacy or saving lives. This is a complex question that cannot be addressed without an understanding of the different operational models, technological parameters and governance mechanisms being employed in each case. And, it presents a false choice where the true potential lies in how these imperatives are reconciled. COVID-19 has made the need for safe, sustainable public-private data sharing models clear and raised the stakes around the need for cooperation, investment and standard-setting to maintain trust in data sharing initiatives.

## DATA FOR COVID-19 RESPONSE

A principal type of data being used is location or mobility data, accessed from different platforms. The range of data sources for COVID-19 response include: GPS, mobile cell towers, Wi-Fi networks, Bluetooth connections, surveillance video, social media feeds, credit card records, wearables, and many other devices and apps.

Key uses of this data include: flows mapping, proximity and contact tracing, misinformation / the infodemic, self-assessment and symptom monitoring, surveillance and quarantine enforcement, and epidemiological modelling.<sup>4</sup>

# Our approach

The learning series was structured around the priorities and needs identified by participants through the preparation phase and the first session.



# Who Participated?

## ALL SESSIONS: SUMMARY OF PARTICIPANTS

**32**

Participants

**10** Countries

**23** Organizations

Intermediary (14)

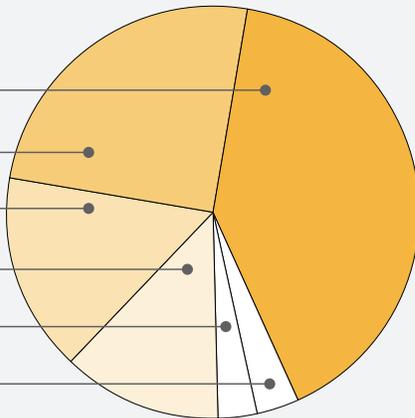
Data User (8)

Data Holder (5)

Researcher (4)

Donor (1)

Civil Society (1)



## SESSION 2

**12**

Track A Participants

**12**

Track B Participants

High engagement and openness in discussion across both tracks

Track B had slightly more diverse representation of stakeholder types

Track A was relatively heavy on intermediaries

## SESSION 3

**13**

Track A Participants

**14**

Track B Participants

High levels of engagement and openness across both tracks

Track B again had more diverse representation across stakeholder types

# Learning Themes Defined



The complexity of **navigating the data sharing ecosystem** and the range of actors, institutions, and needs at play in data sharing initiatives.



The need to improve **communication and transparency** to increase understanding of the value of data sharing and to promote public trust.



The need for **common standards, frameworks and policies** to guide and facilitate data sharing efforts and the potential role of **common infrastructure or platforms** to enable data sharing at scale.



The **human dimensions** of data sharing initiatives, including thinking about the range of actors involved, their perspectives, capacities, roles and interests, as well as the impacts on people and all the elements that go into safeguarding against harms.



The challenge of **sustainability**, specifically financial and institutional approaches that provide fair value to all actors.



The value of bringing stakeholders together for **knowledge exchange and joint learning** around how to build initiatives intentionally and responsibly.

Based on these themes, we structured discussions around **two topics:**

**Data sharing ecosystem—  
Policies, protocols and standards**

**People matter—  
Capacities, trust and ethics**

# Data Sharing Ecosystem—Policies, Protocols, and Standards

Participants mapped common pain points that arise in establishing and implementing data sharing partnerships.

## URGENCY IS DRIVING INNOVATION, BUT WITHOUT FIRM FOUNDATIONS.

Partnerships have been established quickly to respond to Covid-19 but they are mostly ad hoc. They cannot be sustained without establishing a firmer policy framework, clear and replicable data sharing protocols, and financially sustainable models that return fair value to all actors, and practitioners have difficulty finding rapid guidance and advice that meets their needs.

## CONCERNS AROUND PRIVACY STEM FROM POOR COMMUNICATION AND LACK OF TRUST.

Covid-19 has raised the visibility of data sharing initiatives with more media coverage than ever before. Poor communication and relatively low data literacy has led to misunderstandings about objectives and safeguards.

## WHAT DOES SUSTAINABILITY LOOK LIKE?

Finding a sustainable model is a recurring challenge, particularly making the transition away from a philanthropic arrangement. Understanding and defining value has been a major challenge particularly on the demand side where the returns are not well-understood.

## WHERE TO START AND HOW TO SEQUENCE ACTIVITIES?

There is a complex mix of technical, governance and partnership questions to address when embarking on a data sharing initiative. How-to guides grounded in good practices are valuable but they remain difficult to find and customize. Diagnostics and other support that can help identify context-specific steps based on available capacity, governance arrangements and technology are needed.

## TECHNICAL VS GOVERNANCE SUPPORT.

The landscape of guidance, tools and support for technical needs is well-developed. On the policy and governance side, common protocols, templates and human capacity needs still require substantial attention.

We prioritized sustainability for a deeper dive

# Sustainability and Common Infrastructure

## WHAT DOES SUSTAINABILITY LOOK LIKE?

**Fair and scalable business models**—All parties get fair value from data sharing - defined financially or in other terms. This depends particularly on a demand-side understanding of value in terms of what problem the data will help solve or what service it can help deliver.

**Technical capacity and IT infrastructure**—Sustainable skills and IT infrastructure supports data access, transfer, storage, processing and analysis.

**Data discoverability**—A platform or market exposes what data are available from a range of holders

**Robust and agile governance**—An established legal and policy framework supports data sharing, along with ethics mechanisms and guidelines. It is robust enough to contend with different use cases, which means moving beyond project-specific arrangements.

**Clear protocols**—These depend on the overarching governance arrangements and provide clear rules for data handling, roles and responsibilities of key actors, limitations and safeguards.

## A ROLE FOR COMMON INFRASTRUCTURE?

Common infrastructure, also referred to as platforms or frameworks, are being developed as one pathway toward sustainability. They aim to leverage economies of scale by encouraging many holders and users into a common platform that addresses the many dimensions of sustainability. Examples include:

- World Bank Development Data Partnership
- Global Pulse - Global Data Access Framework
- The Commons Project
- UN Global Platform
- EU data strategy including a single data market

Consensus on the need and broad vision for a common infrastructure, but how do we get there from here? Practitioners working at country-level urgently **need solutions that will deliver in the short-term while designing for the long-term.**

# People Matter— Capacities, Trust, and Ethics

Participants examined the importance of human factors in determining how and how well data sharing initiatives work and whether they are sustained.

## HOW TO BUILD TRUST?

**Clarity is essential.** People need clear protocols around purpose, roles and responsibilities for data access and handling, and limitations.

**Terminology matters.** Establishing a common understanding of terminology and implications for data standards and methods is essential.

**Ethics need to be considered up front** and then addressed by defining limitations and safeguards around use.

**Transparency and effective communication matter** both among partners and with the public.

## WHAT DO PEOPLE NEED?

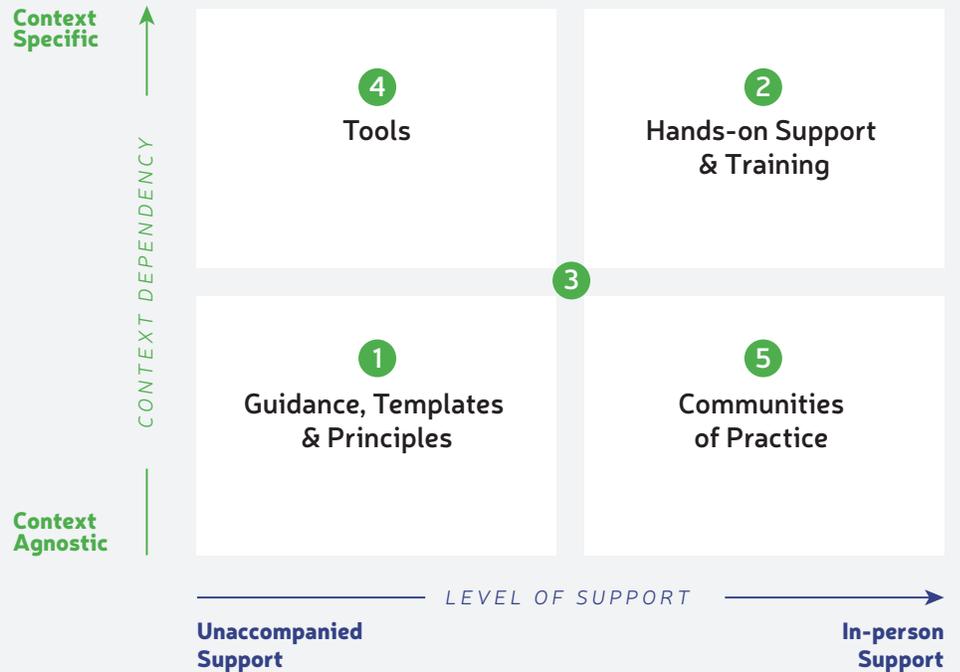
**Tools** based in good practice and diagnostics to support customization to context-specific challenges and needs

**Guidance, templates and principles** to better define standards and protocols for strengthening interoperability, accountability and transparency

**Hands-on support and training** particularly to support the role of data stewards (on demand and supply sides) and to increase data users' understanding and capacity to engage in data sharing partnerships

**Communities of practice** and systems for ongoing peer and institutional exchange

# Getting people what they need



## CONSIDERATIONS AND GAPS FOR TOOLS, GUIDANCE AND SUPPORT:

- 1 Differentiating resources:** Needs for guidance, tools and other support vary depending on the role and positioning of different actors in the data sharing ecosystem, with technical users having more context-specific needs (e.g. around specific data use) and management users having more context-agnostic needs (e.g. around justifications).
- 2 Finding hands-on advice:** Hands-on support and advice are needed to navigate the landscape of tools and guidance, identify what to use depending on context, maturity and existing resources, and customize it.
- 3 Bolstering governance support:** There are significant gaps in governance support compared with technical resources and intermediaries.
- 4 Assessing use:** There is a disconnect between practitioners' expressed need for tools and guidance and the vast landscape of existing tools, suggesting that there is a need to better understand which tools are actually used and why.
- 5 Facilitating exchange:** There is a value in ongoing exchange of knowledge and good practices to align value propositions and avoid duplication.

# The time is now

The interest and need to enter into data sharing partnerships between public and private actors has increased dramatically in response to Covid-19. Many users in government institutions who did not see the value in the early days of the pandemic, are now seeking out these partnerships. **It is critical to seize this momentum and scale data sharing safely and responsibly.**

While challenges exist on all sides of data sharing partnerships, the discussions in this series point to **greater challenges on the public sector data user side and a failure to adequately consider the user perspective, which will continue to hold back progress if left unaddressed.**

## RECOMMENDATION THEMES



**People matter  
and they need help**



**Sustainability  
depends on fair value  
for all sides**



**A role for common  
infrastructure**

## RECOMMENDATION:

# People matter and they need help



### **Develop support and training with specific user needs in mind.**

There is a complex landscape of actors on the user side of data sharing partnerships from managers and decision-makers to analysts. Training and support offerings should be tailored to the different needs of these actors, including technical skills, negotiation skills, governance and ethical issues, and interpreting, understanding and communicating the insights derived from data.

### **Foster more matchmaking to help users find what they need.**

Matchmaking across the ecosystem should raise the visibility of user challenges, and connect them to relevant intermediaries and other providers of support.

### **Clarify terminology and definitions across the technical and governance dimensions**

of data sharing to build trust. A general glossary of terms that can be customized to meet specific needs would be valuable and could also support clearer and more transparent communication with the public.

### **Strengthen support around governance issues.**

While many intermediaries specialize in the technical aspects of data access, processing and analysis, there are fewer equivalent supports on the governance side .

### **Assess the use of tools, guidance and other materials.**

There is a disconnect between the expressed needs for support among less experienced practitioners and the many tools, guidance materials, and principles that exist. It would be valuable to better understand if these tools are reaching their intended audiences and meeting their goals.

## RECOMMENDATION:

# Toward sustainability: Getting a better handle on valuation



**Build user understanding of the value of data in terms of what it will enable them to do.** To establish sustainable business models, users need to understand the value of data in terms of how it can inform their decisions and what it will enable them to deliver. Packaging data as a service by communicating what it will enable decision-makers to do, rather than just describing the indicators, can contribute to building this understanding.

**Develop data valuation methodologies.** Develop and build consensus on robust data valuation methodologies that can be applied across different contexts in order to help practitioners communicate the value of data sharing partnerships and to support the development of fair compensation models between data providers and data users.

**Develop robust impact assessment methodologies across the project life cycle.** Impact assessment methodologies can contribute to a better understanding of whether to go forward with a data sharing project, what models work well under what circumstances, what are the critical factors for success and what are the barriers to translating insights into policy or action.

## RECOMMENDATION:

# Build common infrastructures with the user in mind



There is wide support for an infrastructure, framework or platform that can support data sharing at scale by providing a common legal foundation, parameters for responsible data use, common standards and protocols for data access and exchange, and a pathway to sustainable business models. Several actors are working towards this, but it remains a long-term aspiration and therefore is not addressing the acute needs of practitioners. As these efforts proceed, **it is critical to ensure that they are informed by user needs and work to deliver against them** in the short-term while designing for the long-term.

**Stay connected to real user problems.** Instead of starting from the infrastructure, platform or framework itself, begin with a list of services that data sharing will allow the user to deliver, or decisions it will enable them to make.

---

**Assess past experience.** If standards for data access and exchange are a central value proposition of the development of the infrastructure or platform, make sure to understand why other standards have or have not worked and what was learned from their development and implementation.

---

**Rigorously examine incentives on all sides.** Will the infrastructure enable public sector data users to solve a problem or deliver a service more efficiently or effectively? Do private data holders see the potential for more than philanthropic returns from participating? Digging into these incentives will help to surface persistent barriers and ways to address them.

# Conclusion

The COVID-19 pandemic has raised the stakes around public-private data sharing leading to partnerships that would not have been possible even a few months ago. While this experience will likely pave the way for more public private collaboration around data in the future, the collective experience of the diverse participants in this series points to persistent challenges and needs especially for users that are holding back progress. Focusing on user perspectives and needs will be critical to driving demand for data sharing in the future and building responsible and sustainable partnerships.

The Global Partnership for Sustainable Development Data will build on the recommendations from this series to shape our learning and knowledge sharing strategy, a key component of which will be to broker training and other support with a particular focus on data users. We will also keep facilitating collaborative learning. Participants valued the insights that came from the diversity of practitioner perspectives as well as having the opportunity to reflect alongside their ongoing projects. We are committed to building a vibrant data ecosystem fostering innovation and learning from experience.

# Endnotes

1 UN Global Pulse and partners, Global Data Access Framework, [https://www.itu.int/en/ITU-T/extcoop/ai-data-commons/Documents/Opportunities%20of%20data%20sharing%20and%20the%20Global%20Data%20Access%20Framework\\_Data%20Commons\\_UN%20Global%20Pulse.pdf](https://www.itu.int/en/ITU-T/extcoop/ai-data-commons/Documents/Opportunities%20of%20data%20sharing%20and%20the%20Global%20Data%20Access%20Framework_Data%20Commons_UN%20Global%20Pulse.pdf); The World Bank and IDB, The Development Data Partnership, <https://datapartnership.org>; UN Stats, The Global Platform, <https://unstats.un.org/unsd/statcom/51st-session/documents/>

2 The Open Algorithms Project, OPAL, <https://www.opalproject.org/>; CASD, Secure Data Hub, <https://www.casd.eu/en/>; Flowminder and DIAL, FlowKit, <https://web.flowminder.org/publications/flowkit-unlocking-the-power-of-mobile-data-for-humanitarian-and-development-purposes>; The GovLab, DataCollaboratives, <https://datacollaboratives.org/>; The Commons Project, <https://www.thecommonsproject.org/>

3 The GovLab, Data Stewards Network, <https://www.thegovlab.org/project-data-stewards-network.html>; GSMA, Big Data for Social Good, <https://www.gsma.com/betterfuture/aiforimpact>

4 Andrew J. Zahuranec and Stefaan G. Verhulst, Mapping How Data Can Help Address COVID-19, Data & Policy, <https://medium.com/data-policy/mapping-how-data-can-help-address-covid19-a7be2e631aec> and Nuria Oliver, et al, Mobile Phone Data for Information Public Health Actions Across the COVID-19 Pandemic Life Cycle, <https://advances.sciencemag.org/content/6/23/eabc0764/tab-pdf>.