



Global  
Partnership  
for Sustainable  
Development Data

ATHENA  
INFONOMICS

# Microsoft Open Data Campaign

## Facilitating data collaborations to overcome the 'data divide'

### Summary

This case study was developed as part of *Effective Data Sharing: Beyond Platforms*, a study conducted by Athena Infonomics and commissioned by the Global Partnership for Sustainable Development Data. This project was made possible by a grant from Google.org.

The data revolution has led to tremendous growth in the amount of data available, and the growth of new data computational techniques, such as AI and machine learning, have improved the potential use of this data. However, access to data and the capacity to harness new technologies remains concentrated in a limited number of organisations. This divide could further exacerbate inequities across geographies and societies.

In April 2020, Microsoft launched the Open Data Campaign ('initiative') to promote data sharing and collaboration so that organisations of any size can avail the benefit of new data and technologies. One of the key aims of the initiative is to overcome the 'data divide'. The initiative seeks to facilitate data collaborations that address significant social, economic, and environmental challenges. The initiative recognises that the value of data lies in how it is analysed and how the insights generated are used. The initiative seeks to facilitate partnerships between different organisations so that together they can unlock the value of data.



Source: Microsoft Open Data Campaign

## Facts and figures

**Founded:** 2020

**Sector:** Cross-Sectoral

**Typology by use:** Capacity Building and Knowledge Transfer

**Geography:** Global

**Governance Structure:** Data Collaboration

**Number of Data Collaborations:** 20 data collaborations facilitated by 2022

## Key Challenge(s) that the initiative was trying to solve

- Enable access to data and its benefits to organisations of all sizes; develop the ability to share data across organisational boundaries.
- Building essential assets that will make data sharing easier, including the required tools, frameworks, and templates.
- Launching new data collaborations that address significant social, economic, and environmental challenges.

## How are they solving the problem?

The initiative's primary aims were to facilitate data collaborations and to create common resources that can facilitate data sharing across organisations. The initiative adopted five principles that will guide its data collaboration approach. These include: **Openness** of data, **Usability** of data, **Empowering** organisations to generate value from their data, **Secure** sharing of data, and Protecting **Privacy** of data by removing personally identifiable information.

**“Often the first questions, when you go into a data collaboration, are what data is to be shared and how are we going to share it, but really some of the biggest hurdles were these topics on the front end. Before you even get to the technology, the issue of leadership, understanding what the opportunity is, the talent needed for it and how collaboration would be governed. These are critical to addressing before coming to tech and data.”**

The initiative's early experience suggested that helping organisations build an internal roadmap to share and open their data was crucial to promoting data-sharing collaborations. Before arriving at the data question, questions of trust, leadership, opportunity, and governance needed to be addressed. The initiative created an [Open Data Social Impact Framework](#), which would help organisations better use data. Along with the Open Data Institute, the initiative made use

of a [Data Ecosystem Mapping Tool](#) to help each stakeholder identify the value of sharing data. Thus, the initiative adopted the role of a collaboration facilitator which builds the tools and capacities needed to build data partnerships. Once a collaboration had been established, the initiative could further act as a data partner or a technology partner, depending on the nature of the collaboration. For example – Microsoft shared its sustainability data to [OS-Climate Platform](#); it is an initiative by Allianz, Amazon, Microsoft and S&P Global that uses open-source analytics and open data to assist companies in modeling their risk to climate change. Whereas it provided technical support in the form of cloud computing in the [Concentrated Animal Feeding Operations](#) Project; a project to map industrial poultry operations with deep learning and aerial imagery.

## What were the key considerations in designing the initiative?

**1. Making Data Usable:** The initiative recognised that not just access to data, but access to usable data was a key component of overcoming the data divide. Inconsistent data collection and variance in data formats and documentation hinder data sharing. Through data collaborations, the initiative sought to work with organisations to make their existing data usable for the specific needs of the collaboration.

**2. Creating common tools and resources to ease data sharing:** From the outset, the initiative was keen to develop a knowledge repository that would create tools and resources to benefit data sharing more generally. The initiative tied up with the Open Data Institute and The Gov Lab at NYU to develop resources, courses and networks that could help organisations and leaders navigate the transition to a more open and collaborative data approach and help them develop the internal governance structures and capacities needed for the same.

**“Many organizations want to do more around open data and data sharing, but when it comes to the practical aspects of how to do it, they often don’t know where to start.”**

The initiative has worked with its knowledge partners to develop resources which will help provide data stewardship guidance for public sector and private sector actors. This includes the creation of a new [Data Stewardship Academy](#); an [Open Cities project](#) to build community and share insights among cities using data to drive change; a [Peer Learning Network](#) which will act as a community of practice and learning.

**3. A 'one-size-fits-all' approach would not work:** The initiative facilitates different types of data collaborations with a range of public and private sector actors. This has meant that the role of the initiative and the manner in which the collaboration came together varied widely. Across the different collaborations, the initiative has encouraged the use of methodologies that can help guide the structure and manner of collaboration. An example is the [Data Ecosystem Mapping tool](#) which helps identify different data actors in the ecosystem and the value that can be exchanged through the collaboration. In a recent data collaboration in New Zealand, the data ecosystem mapping methodology was used to bring a number of stakeholders, including indigenous actors, together. This methodology was used to develop trust amongst stakeholders and build a participatory governance model. The manner and structure of the data partnership are underlaid by trust between the partners and are co-defined by the collaborators. This can take different approaches and the role of the initiative can vary from a collaboration facilitator to a technology partner.

## Financial Sustainability of the initiative

There is no common funding approach for the data collaborations. Each has a different business model, with some identifying revenue structures for the collaboration. For example – OS-Climate platform which uses data science to produce climate sustainable investment models has an annual membership fee model for revenue generation.

## Lessons Learned

- **Organisational leadership and commitment are crucial to unlock the value of data sharing and insights.** Leadership buy-in, identification of a clear use-case, and assembling the necessary talent is crucial to creating value from data and using it to improve outcomes. Data sharing initiatives include not just a technology component but require cultural, strategic, and operational shifts. Strong champions within an organisation are needed to overcome any internal barriers and to develop a clear human resourcing and governance approach to sharing data. Data sharing is about more than just data.
- **Clarifying the problem statement that the data collaboration is trying to address is vital for success.** One of the key drivers of success for data collaborations has been a clear value identification and defining what each stakeholder is getting in return for sharing the data. This is important not just for the other partners but also in generating internal organisational buy-in for the collaboration.
- **Flexibility and change must be built into data collaborations.** The initiatives' initial collaborations highlighted the importance of the ability to pivot and change direction when needed. An early use case that looked at air quality was repurposed to measure the impact of the COVID-19 lockdown. While this was not the original intent of the data collaboration, the data streams had already been put in place and the data infrastructure was used to generate insights that were relevant to the COVID crisis. Trust between data partners allows for such flexibility.

The Microsoft Open Data Campaign acts both as a facilitator and knowledge repository for promoting open data and data collaborations. The initiative's work highlights the importance of building organisational capacity and common resources to unlock the value of sharing data. The collaborations facilitated have unlocked analytical insights that help large-scale societal challenges such as [climate change](#) and [health care research](#). Continued focus on the societal impact of data sharing, while working to build capacities for data sharing across varied organisations and geographies is important.

## Sources

Stakeholder Interviews

Yokoyama, [Closing the data divide: the need for open data](#), 2020

Microsoft Open Data Campaign, [Open Data for Social Impact Framework](#), 2021

Wiggers, [Microsoft provides progress update on Open Data Campaign](#), 2021

Hardinges, et al., [Microsoft and the ODI – an update and reflections from the second year of our partnership](#), 2022

Yokoyama, [Building on a year of open data: progress and promise](#), 2021