Spotlight

Unlocking data for climate action requires trusted marketplaces that promote transparency and control. Data spaces offer a promising solution.

Joint Learning Network on Climate Action

digital impact alliance

In 2024, the northern hemisphere recorded the hottest summer overall, the hottest day, and the hottest ever month of August. That same month - August 2024 - this warming fueled droughts in Italy and intensified typhoons that devastated parts of the Philippines, Taiwan, and China. The following month, new research calculated that warming is costing the global economy billions of dollars: an increase in extreme heat and severe drought costs about 0.2% of a country's GDP.

These are only the latest stories and statistics that illustrate the growing costs of climate change - data points that have emerged in the short time since we published our second Spotlight on unlocking climate data with open transaction networks. This third paper in the series continues the work of the Joint Learning Network on Unlocking Data for Climate Action (Climate Data JLN), which was launched under the ITU's multistakeholder Green Digital Action initiative. The Climate Data JLN brings together experts in climate action, data exchange, and digital public infrastructure to explore new models for data sharing to help frontline actors, namely city governments in low-resourced countries, address people's urgent needs and build resiliency in the face of this global crisis.

This multi-disciplinary network identified multiple promising models to explore in the context of unlocking data for climate action. Launched at COP28 in Dubai, the Climate Data JLN will deliver recommendations in time for COP29 in Baku.

This Spotlight paper examines the third of these models: data spaces. Through examination of data spaces in action, the paper analyzes the key elements that render them more or less applicable to specific climate-related data sets.

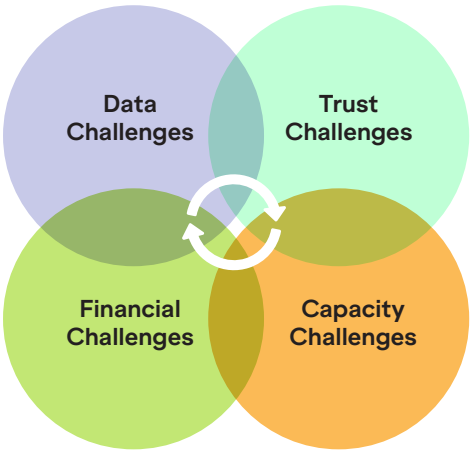
Data spaces are relatively new and mostly conceptual, with only a handful of implementations in process and concentrated in a few geographic areas. While this model requires extensive up-front work to agree upon governance and technical standards, the result is an approach that overcomes trust and financing issues by maintaining data sovereignty and creating a marketplace for data exchange.

Much of this work was inspired by the multistakeholder Green Digital Action (GDA) initiative, convened by the International Telecommunications Union with participation from 40+ organizations.

Following our engagement on the GDA track at COP28, the Digital Impact Alliance launched the Joint Learning Network on Unlocking Data for Climate Action (Climate Data JLN.)

Green Digital Action aims to enhance collaboration, fast-track industry-wide commitments to addressing climate challenges, and put digital solutions at the forefront of climate action.

Challenges to the use of data for climate action



Several challenges impede the effective use of data for climate action. These challenges, identified through early discussions with the Climate Data JLN, were described in our first Spotlight papers on data trusts and open transaction networks. They were broadly categorized into issues related to data availability and quality, trust, financing, and capacity.

Challenges related to data availability and quality

- Data may be incomplete, inaccurate, or out-of-date.
- Climate-relevant data is often fragmented and scattered across various non-interoperable platforms, leading to inefficiencies
- Current technology stacks are not always interoperable, meaning that data cannot be easily shared between systems on a real-time basis.
- When data is collected at a global, regional, or even national level, the utility of that data at the local level may be limited due the lack of granularity or comparability.

Challenges related to lack of trust:

- Many countries lack applicable or robust legal and regulatory frameworks for issues such as intellectual property rights and data transfer restrictions, and where these frameworks exist, cross-border border disparities can provide additional obstacles.
- Current data owners may be hesitant to share data due to data privacy concerns, their inability to negotiate proper terms and conditions, or the lack of oversight and dispute mechanisms.

Challenges related to financing:

- Data owners often do not want to share data without a return on investment, which can result in paywalls that are unaffordable for many frontline responders who operate in low-resource environments.
- Governments, particularly sub-national governments, often do not have continuous budget streams allocated to the on-going purchase of data.
- The high transaction costs associated with discovering and fulfilling climate-aligned goods and services pose significant barriers.

Challenges related to capacity:

- Frontline, sub-national governments often do not know what digital platforms and data solutions exist. If they are aware, they often lack access to technology, a culture of data-driven decision-making, and/or staff trained to analyze data and translate it into information relevant to solutions.
- Inefficient public systems for permits, applications, and subsidy delivery exacerbate these challenges, making it difficult to implement effective climate action at scale.
- Consequently, there is a significant gap in fast, efficient, and secure data-sharing solutions critical for climate action. This gap results in inactivity and the underutilization of vast data resources.
- To bridge these gaps and overcome some of these challenges, data spaces present one promising solution. As we will see in the following case studies, data spaces are particularly well-suited for overcoming challenges related to trust, as they enable data owners to maintain control over their own data, and financing, as they allow for data owners to set their own price for sharing data, within the bounds of agreed upon rules.

How data spaces can address some of these challenges

What is a data space? In fact, as this concept is quite new, there is no one set definition. That being said, most would agree that data spaces are a model of data sharing that enables the reliable exchange of data while retaining sovereignty, and ensuring trust and security under a set of mutually agreed rules (AD4GD).

The 2020 European strategy for data established the concept of data spaces in order to overcome some of the key factors keeping data locked away: namely, that public and private sector entities often lack incentives to incur the costs associated with data sharing and don't have the resources to navigate the various legal and technical barriers required to minimize risk when sharing data. The key features of data spaces aim to overcome these barriers. They include:

Sovereign data control: Data owners maintain control of their data and the conditions for data sharing - establishing rules about who can see their data, and for how long. This feature differentiates data spaces from data trusts - where data is pool and held by the central governance body.

Monetization: Each member can set their own monetization policies and fees (or choose to offer data for free), thereby creating a market for incentivized data sharing.

Centralized governance: Ideally, there is a neutral and trusted actor who enforces rules regarding membership, data quality, and community interaction.

Decentralized data transfer: Data is shared through APIs that comply with set standards. There is no one central platform or data repository. This allows for federation, and presents the possibility of allowing one data space to link to others that are focused on different sectors or geographic locations.

It follows that data spaces are unique from data trusts or other data sharing mechanisms such as data lakes, since data is never pooled. Rather, it is shared through a bilateral connector which enables both peer-to-peer, peer-to-many, and many-to-many data transfer.

In a 2023 Data Diagnostic for Kerala, the World Bank identified data spaces as a promising option for the Indian state because climate "resilience is the classic boundaryless modern sector." Since those actors at the frontlines of climate change - namely, local governments and community organizations - require data from across traditional sectors such as water, energy, transportation, and agriculture, they need data sharing options which are unencumbered by traditional organizational, national, and sectoral boundaries. With proper governance mechanisms and standards in place, data spaces can play this role.

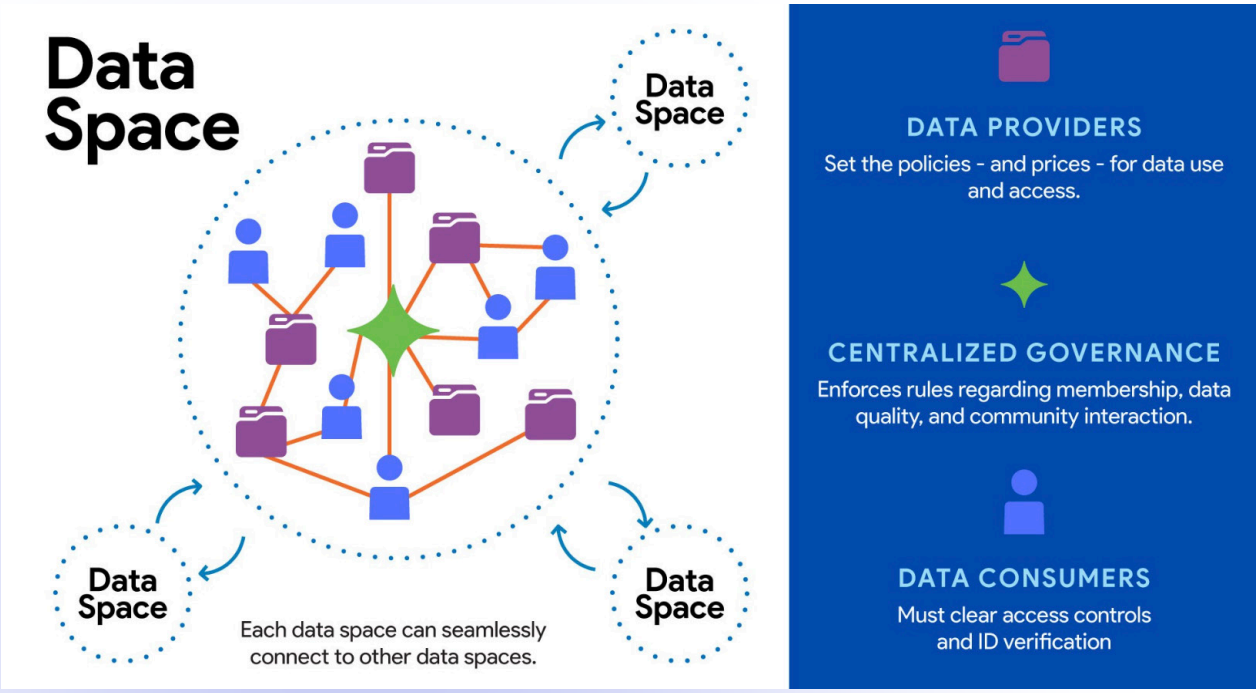


Figure 1: Data space visual explainer

Data space evolution

Data spaces as currently conceptualized are largely a European project, although they echo and strengthen efforts emerging from other regions to improve cross-sector and cross-border data sharing, such as Japan's efforts to promote data free flow with trust (DFFT). The European strategy for data declared the intention to create a single European data space - in other words, a single market for data - that will boost growth within the framework of EU law and regulations. To move toward this unified data market, the strategy is to first create 14 Common European data spaces.

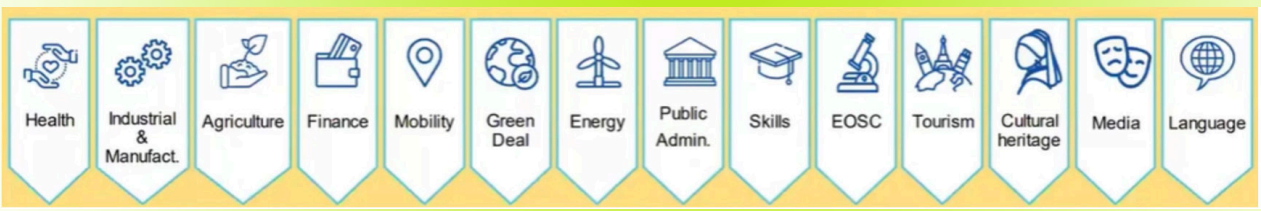


Figure 2: Image from All Data for Green Deal (AD4GD)

Each of these 14 data spaces is at a different stage of development, but all have the same intention: to overcome legal and technical barriers to data sharing by building the technical, legal, and governance infrastructure required for interoperability and reuse. The Green Deal Dataspace (GDDS) will be of particular relevance to climate data; however, it is not yet under development. The GDDS will support the Europe Green Deal, the EC's ambitious roadmap to become the world's first resource-efficient and competitive economy, achieving net-zero greenhouse gas emissions by 2050. The Green Deal underscores the pivotal role that data will play in this transition, and advocates for accessible and interoperable data to be at the heart of Green Deal innovation. To this end, the GDDS will accelerate the sharing of high value and high quality datasets for key climate objectives including biodiversity preservation, zero pollution, the circular economy, deforestation reduction, and environmental compliance.

The GDDS, like many of the other 14 data spaces, has yet to move to implementation because of the upfront work that needs to be done to build agreement on the standards and trust frameworks that will underlie the data space. Multiple organizations are helping to drive this collaboration, including the International Data Spaces Association (IDSA), Gaia-X, Fiware, the Data Spaces Support Centre (DSSC) and the Data Spaces Business Alliance (DSBA). These

organizations work together to keep the specifications and standards aligned and to avoid multiple overlapping or competing specifications.

Recognizing that data spaces are in the early stages of development, (one expert we spoke to described the current data space landscape as the “Wild West”) there is sufficient testing underway to explore how this approach can unlock relevant data for climate action. Particularly as global players start to converge around a shared set of standards and practices, data spaces should mature and become more widely adopted. The following section will dive into two examples that are relatively far along in their implementation and which show how data spaces will eventually translate into improved climate resilience for people and communities.

Data spaces in practice

Mobility Data Space

Overview: The Mobility Data Space (MDS) is the German effort to push forward with a practical implementation of the European Mobility Dataspace mentioned above, and is profiled here due to its relatively advanced implementation. MDS creates a virtual marketplace for the exchange of mobility-relevant data. It brings together companies, organizations, and institutions, connecting those who need data for innovative mobility solutions with those who have such data and wish to monetize it.

By fostering a data economy, the MDS aims to create an environment where data assets can be exchanged efficiently and securely. In addition, through the combination of different datasets and providing new possibilities of collaboration between different stakeholders, the MDS allows for innovative new business opportunities Ultimately, this data will be used by cities and other actors to increase road safety, improve public transport and multimodal mobility, and accelerate the use of electric vehicles.

Technology: The architecture is closely associated with Gaia-X and IDSA and emphasizes secure data exchange through bilateral or multilateral means rather than a centralized platform. Data purchasers can see all available data through a catalog, and then connect directly with data providers to acquire data for a set purpose and price, which is then shared through a decentralized data exchange protocol. The decentralized, or federated, architecture allows for the sovereign sharing of data, as owners can define who can use their data, for which purposes, and under what conditions.

Partners and governance: The MDS was developed in 2019 in coordination with more than 200 German mobility stakeholders, steered by the acatech Foundation. In 2021, the acatech Foundation helped to establish a non-profit holding structure that serves as a neutral governance body that can moderate conflicts, enforce rules, and reduce transaction costs. It is noteworthy that the MDS initially

attempted to operate without a central body, but this proved unsuccessful, highlighting the importance of oversight.

Regulatory environment: European Union rules and regulations significantly influence the MDS, which is actively providing insights to the European Commission. The need to align with the developing landscape in Europe presents both opportunities and challenges, as increasing collaboration with other data spaces is fostering the development of common standards. For instance, the MDS first adopted the International Data Spaces Association (IDSA) reference architecture, and then later added the Eclipse Dataspace Components (EDC) connector as it became available. The EDC provides a comprehensive framework (concept, architecture, code, and samples) that builds on the IDSA protocol and the Gaia-X trust framework. As these policy and standard-setting conversations continue, the MDS is able to play a role in shaping broader policy discussions, particularly as they relate to the ongoing development of the European Mobility Data Space.

Status: The MDS has established both the technical and governance infrastructure, and has progressed further than most European data spaces. 190 partners have signed on as participants, including public sector actors, large companies, and small startups. At the same time, the MDS has yet to move past bilateral data sharing use cases, and thus it has yet to exploit the full potential of the data space. These bilateral partnerships are complex, and take on average one to three years to establish since trust, legal, and business concerns need to be addressed. Ultimately, establishing these partnerships within the context of the data space will help these bilateral connections to evolve into a larger network. As more use cases emerge, the potential for complex data sharing arrangements and the development of common standards will continue to grow within the evolving data space landscape.

Sectoral Data Spaces in West Africa

Overview: The Data Governance in Africa - Team Europe Action is a joint project involving the European Commission, the African Union Commission, Germany, France, Belgium, Estonia and Finland focused on three main components: data regulatory frameworks, data use cases, and data infrastructure. Under the data use cases component, GIZ is piloting an approach with sectoral data spaces, designed to break silos and create local value out of data in areas such as climate and agriculture. Sectoral data spaces aim to produce local value with strong protection of data sovereignty.

GIZ and its partners selected a sectoral data space approach only after exploring several other data sharing options because it was the only option that, by design, can overcome many of the data flow and usage barriers faced in West African countries. To start with, data space technology offers the possibility of interconnecting already existing national and international databases and data platforms. Furthermore, a data space integrates public and private data and creates the right incentives for each stakeholder based on business models and data sharing agreements. Since data spaces are open source and modular, they offer the possibility for constant growth by integrating new utilities such as data marketplaces for business-to-business exchange and AI-on-demand. Finally, data spaces are community-based and the governance is assured by its members.

Technology: The West Africa approach builds from the EC Common Data Space policy, and benefited from project financing in developing open-source technology. GIZ is using a brownfield approach by upgrading and interconnecting existing data platforms. While the technology is already freely available online, the majority of the work is dedicated to the governance of the data space - creating the incentives, sharing agreements, usage rights, and many other aspects designed to create fairness and trust among the participants.

Partners and governance: This initiative has kicked off in Senegal and Ivory Coast. In each country, the national government is the main partner, together with a large number of stakeholders from the private sector, civil society, and academia. All the results are co-designed by a multistakeholder Core Team.

In order to ensure national sovereignty, GIZ plays a support role, conducting training and technological transfer to local stakeholders and startups who then take the lead in developing the data spaces. For each data space, the work begins by developing a roadmap with all local partners. Through this collaboration, the national government and local stakeholders determine the governance model.

Regulatory environment: The project is working with the National Data Protection Authorities in each country to create a regulatory sandbox for a specific period of time, approximately two to three years. The sandbox approach allows a smaller group of participants to form a data cooperative and test different rules such as access rights and licenses in a secure and supervised environment.

Status: The agricultural data space in Senegal is the furthest along. Senegal's existing national AgriData platform is being upgraded to operate as a data space which will provide targeted advice to small farmers and allow the government to deliver its services more efficiently. The initial work to establish partners and collaboration has begun in the climate sector and in Ivory Coast.

Making data spaces work for climate action

Drawing from these examples of data spaces in action, the Joint Learning Network identified key areas that need to be further explored in order to make data spaces work for climate action. Each of these areas are ultimately a matter of effective governance. While data spaces seem at first glance to be highly technical, the technology is actually relatively straightforward. It is the governance that is critical to maintaining the underlying trust that drives the uptake and effectiveness of the data space, as the MDS learned when it attempted to first establish a fully decentralized network. The fact that there are many more data spaces conceptualized today than there are implemented is a result of the need to establish governance, and by extension, trust, through extensive consultation before any data sharing actually occurs. Once implementation starts and sharing is underway, the governance body and the participants need to continually work to establish and enforce rules and and norms that drive context-specific incentives for data sharing, navigate tensions associated with monetization, and prevent abuse.

Understanding context-specific incentives for data sharing

In different markets, different actors hold data, which changes the incentives for unlocking that data. For example, in the U.S., large corporations tend to hold most of the data. On the other hand, as we see in the European push for data spaces, there are many public and private sector actors that hold smaller amounts of data, and they are often encouraged through regulation to share their data. In many countries in Africa, including Ivory Coast and Senegal, there are also many data holders across the public, private, and development sectors, each with their own siloed platforms. In each of these different contexts, data sharing requires a different mix of incentives, ranging from regulation to monetization to community improvement of data shared within the data space. Getting the mix of incentives right will be critical within the context of climate, where the nature of the climate crisis means that relevant data is held by actors across sectors and borders.

Navigating monetization challenges and tensions

Both examples cite the possibility of allowing participants to monetize the data within the data space as one of the key benefits. Monetization is just one incentive for sharing, but it can be a critical way to unlock data held by private sector actors that demand some return on investment for data collection and

cleaning. The need for monetization as an option is underpinned by some of the challenges with the open data movement over the past decades, which has successfully unlocked a range of data but has struggled with quality, access, and sustainability.

However, pricing data is complex. One approach is to price based on the cost to collect and provide the data and add some margin. However, this approach does not take into account the value of that data to the customer, which can range from small (marginal improvement to an AI model, for example) to large (data that is crucial for decision making). Pricing will also need to be dynamic, as the data space can, in theory, increase the number of potential customers and increase volumes. Regardless, data space providers stress that monetization is only one incentive among many, and governance rules should ensure pricing fairness and transparency.

Maintaining trust and preventing abuse

In all of our examples, including the overall European effort to establish a unified data market, it takes years to create the underlying trust that is required for the data space to function in the long-term. To establish and maintain trust, data spaces are using a variety of approaches. First, most data spaces are starting with a sector-based approach, where actors are more likely to know each other and to have clear incentives to share data. Second, some data spaces such as those in Senegal and Ivory Coast are starting with sandboxes and demonstration models. We also see on-going community engagement efforts such as ideation workshops. These efforts can help to work out tensions, such as hesitancy from the public sector to share data from which the private sector can then profit. In a trust-based environment, this value addition can be an incentive: if one shares data, and then others improve upon it and share it back to the data space, it may be worth sharing and paying a small fee in order to access the improved data.

On the other hand, data spaces, like data trusts and open transaction networks, have yet to fully figure out how to address misuse and abuse. It is clear that if any one actor breaks the rules, this can undermine trust and the overall data sharing project. Bad actors can be removed from the network, but it can be hard to define exactly when such enforcement should occur, especially since they may take valuable data away from the network when they leave. Thus, appropriate enforcement mechanisms based on different degrees of overall risk will need to be considered in each context.

Looking forward

Data spaces are clearly a long-term project. By putting in the effort now to establish trust and determine common standards and practices, the actors driving data spaces are putting into place the necessary infrastructure for cross-border data sharing and consent-based AI models - both critical aspects of unlocking data for climate action.

There are signs that data spaces are setting the stage for taking cross-border data sharing to a new level in the coming years. For instance, the Nordic Institute for Interoperability Solutions (NIIS) is currently developing X-Road 8 "Spaceship." X-Road is a trusted national data exchange platform that is deployed by 24 governments and organizations across Latin America, Africa, Europe, and Asia. This evolution will move the open-source platform from X-Road specific protocols to data space protocols, allowing X-Road to connect seamlessly to other data exchange ecosystems that use data space protocols. This shift is seen as critical to allowing X-Road to scale and support cross-border data exchange, as it will no longer be necessary to create custom connections for every different system. All countries that implement X-Road will transition to the new approach, and therefore be aligned with European data exchange initiatives for years to come.

Data spaces also show the potential to support robust, consent-based artificial intelligence (AI) in the coming years. AI models have to be continually trained on high-quality, human-generated data. AI models, whether generative or more traditional machine learning models, can struggle to access new high-quality data where there is a lack of trust. If people and organizations close off their data due to lack of transparency related to who has access and how the data will be used, these models will struggle to remain relevant and useful.

Data spaces are likely to be useful in overcoming these challenges as they curate data sets which are shared based on trust, set rules, and where necessary, monetization. And, as data spaces are interoperable by design, AI models can

be built on top of this data, generating sector-specific models underpinned by a consistent supply of trusted data. Türkiye, for example, is already moving to establish a Public Sector Data Space to support its National Artificial Intelligence Strategy. This will be critical to the climate agenda going forward, as AI is helping with a range of climate priorities including predicting weather, mapping deforestation, and of course, mitigating its own environmental impact.

Unlocking data is no easy feat, particularly at the scale needed to address the constantly increasing and evolving impacts of climate change. To create the trust, safeguards, and incentives necessary to share data at this scale, across traditional sector and country silos, will require massive, sustained effort. The multi-year efforts underway to build the foundations for data spaces whether in Europe, Africa, or Japan are helping to create the foundations for trusted, cross-border data exchange for years to come.

dial.global