Industry Data For Society PARTNERSHIP

Virtual Roundtable on Private Sector Data Sharing

Post-Event Report

May 16, 2024

Table of Contents

Overview

The Industry Data for Society Partnership (IDSP) is an industry-led initiative committed to making private sector data more open and accessible while fostering strong connections between industry and the public sector, researchers, and academia to address societal challenges. The IDSP is composed of a group of companies that share a set of beliefs and commitments around the value and impact of open data and data sharing.

On Tuesday, March 26, 2024, the IDSP hosted a virtual roundtable held under Chatham House Rule with representatives from the private sector as well as IDSP affiliates, The GovLab and the Open Data Institute. The purpose of the convening was to gather insights and recommendations from the participants on how to advance open data and data sharing practices in the private sector. This convening also aimed to identify the needs and gaps to inform the development of a tool or resource to support the private sector and other organizations in their data accessibility efforts.

The agenda included a presentation by Stefaan Verhulst, the Co-Founder, Chief Research and Development Officer, Director, The GovLab; a panel discussion with Casey Weston, Senior Manager of Data for Impact at LinkedIn as moderator and three IDSP members as panelists: Sonia Cooper, Assistant General Counsel at Microsoft; Yiu-Shing Pang, Open Data Manager at UK Power Networks; and Melissa Tallack, Open Data Lead at Northumbrian Water Group.

Breakout sessions were held to identify best practices that participants use to inform their open data and data sharing strategies, as well as to identify the resources or support they consider most beneficial for organizations aiming to enhance their data accessibility efforts.

Key Takeaways

- Private organizations may share data for a variety of reasons, ranging from building trust and transparency to driving innovation and societal benefit.
- Effective strategies for data sharing include purpose-specific approaches, industry alliances, resource pooling, and collaborative awareness-building.
- Common data sharing challenges include a risk-averse culture, regulatory uncertainty, and concerns regarding privacy and security. Stronger data governance and quality assurance mechanisms can help overcome these challenges.

Next Steps

- IDSP members will leverage insights from the roundtable to inform the development of tailored data resource solutions for the private sector, potentially including guides, repositories, toolkits, or communities of practice.

Data Stewardship Pulse: Emerging Trends & Practices

Stefaan Verhulst presented emerging trends and practices around data sharing, providing insights from a global perspective on data sharing initiatives, shedding light on the advantages, obstacles, and facilitators of collaborative data efforts, which included the following.

- recent advancements in global data governance, including notable policies such as the EU Data Act, Digital Services Act, Global Digital Compact, and the US Executive Order on data security and brokers; and how these regulations offer new frameworks for data sharing between businesses and governments, as well as access to platform data for research purposes.
- the significance of non-traditional data sources, particularly privately held data like search, mobility, and satellite data, in addition to synthetic data, in generating unique insights beneficial for public interests.
- the emergence of a "data winter," meaning a backsliding or decline in open data practices and availability, especially in the private sector, despite the potential value for innovation and societal benefit.
- the challenges faced by open data initiatives due to factors such as data monetization, AI, privacy concerns, national security issues, and data hoarding.
- the dwindling attention on open data amidst the increasing focus on AI, emphasizing the need to maintain a balance between the two.

He also mentioned the upcoming UN Summit of the Future in September 2024, which aims to establish key objectives concerning data and digital matters examining the issue on a global scale, particularly focusing on the data divide where those in need often lack access to data.

He hopes that the IDSP can be an example and a platform for responsible and productive data sharing practices that can inform and influence global data governance developments.

Key Themes & Insights

The output below is a consolidation of the panel and breakout session discussions.

Motivations & Incentives

Participating companies shared their motivations and incentives for sharing their data, which include building trust and transparency, enabling innovation, increasing competitive advantage, and driving societal benefit.

Sharing data serves as a catalyst for transparency, bolstering trust among customers and regulators while nurturing innovation both internally and externally.

- One company found that disclosing performance data cultivates transparency and innovation within its sector, empowering stakeholders to make informed decisions and drive enhancements.
- Another participant highlighted that opening up their data invited collaboration, attracting diverse perspectives to enhance innovation efforts.

Sharing data can also lead to cost reduction and enhanced efficiency.

- Another company provided a tangible example, showcasing how open data enhances network efficiency and resilience, benefiting the consumer by passing along the savings and reducing their cost.

Companies can secure a competitive advantage through open data practices.

- Companies who have open data practices are better positioned to comply with regulations and can receive rewards for their data sharing initiatives.
- In regulated sectors, where treating data as presumed open is mandated, companies who have mature open data practices can gain a significant lead over competitors, often up to two years. These companies are regularly rewarded and evaluated for their commitment to data sharing.

Some companies are motivated by their corporate social responsibility goals.

- One participant discussed their company's drive to close the data divide and enable universal access to AI and data development, citing its potential to drive both societal and economic benefits.
- Another participant highlighted the significance of aligning data sharing with environmental, social, and governance (ESG) initiatives, as well as human rights-related endeavors, in contributing to societal and economic progress.

Approaches and Best Practices

Participants discussed their best practices and approaches for opening and sharing their data, which included purpose-driven data sharing, participating in industry alliances, leveraging existing resources, and seeking guidance from organizations with data sharing expertise.

Purpose-Driven Data Sharing

- Companies discussed how they can benefit and build trust partnering with stakeholders, researchers, and regulators to enhance data sharing initiatives and leveraging collective knowledge and resources.
- One example referenced was the New York City (NYC) government's approach to establishing their governance framework for Open Data Coordinators within NYC government. The Open Data Team surveyed data owners and Open Data Coordinators on the resources needed and the challenges they were facing. This early research paired with qualitative research done on New Yorkers' demand for city government data led to the development of NYC's Open Data Strategic Plan. Formulating the data in purpose-driven ways and with specific use cases was informative to drive the strategy forward.
- Another participant discussed how aligning data sharing initiatives to broader organizational goals, such as environmental, social, and governance (ESG) initiatives or human rights-related efforts, made it easier for them to get support for sharing the company's data.

Engagement in Industry Alliances

- Some companies discussed that they actively participate in industry alliances like COVESA, a global alliance focused on the development of open standards and technologies to accelerate innovation for connected vehicle systems or the IDSP, to help them address common challenges in data sharing as well as to foster trust.

Leverage Existing Resources

- Companies shared that they leveraged resources from organizations like the Open Data Institute, The GovLab, and domain-specific groups like the Green Software Foundation. They have also consulted with those organizations to leverage their expertise and exchange knowledge and insights.
- One participant talked about using existing resources such as the ODI's Data Ecosystem Mapping Tool, and another referenced the Open Data Sponsorship

Program by Amazon that enables companies to offset the cost of storing and provides repeatedly requested open datasets which reduced the effort of handling day-to-day data requests.

Awareness Building through Collaborations

- Companies discussed collaborating with research universities and training doctoral students to raise awareness of data sharing best practices and facilitate capacity building within the organization and across industry. They emphasized the importance of metadata for findability and standardization to enhance data usability.
- For example, one participant shared that Airbnb, Booking, Trip Advisor, and Expedia entered into an agreement with Eurostat, the statistical office of the European Union, to publish aggregated tourism statistics across the EU 27, Switzerland, Iceland, Norway and Liechtenstein on a quarterly basis. This approach provides granular insights to the public into short-term rental activity taking place across communities in Europe, while equally balancing privacy dimensions such as data protection for non-professional hosts.

Challenges & Opportunities

Based on the discussions, the table below outlines the challenges and opportunities, with additional context provided by extrapolating the meeting notes.

CHALLENGE	OPPORTUNITY
Culture of risk-averse mindset and resistance to change	• Demonstrate value by creating an environment to encourage a cultural shift toward data sharing through communication, advocacy, and community engagement. Showcase examples of how making data more open or accessible improves results for the private sector and society, or share how using AI can accelerate the process of achieving better outcomes. • Foster transparency and trust by collaborating and partnering with others in industry and the research community, such as joining cross-industry initiatives and alliances to address common data sharing challenges and to establish best practices, partnering with research universities to train doctoral students on

	how to access data, or rewarding organizations for their data sharing efforts.
Regulatory and legal uncertainty	• Establish clear data sharing agreements and protocols, collaborating with legal experts and industry peers to navigate legal complexities and ensure compliance with evolving privacy regulations. • Adapt strategies and processes to comply with evolving policies and regulations and collaborate with regulatory bodies and industry peers to influence policy development.
Privacy and security	• Utilize privacy-enhancing technologies like differential privacy and confidential containers to protect shared data while maintaining usability. • Address granularity and anonymization in data sharing to balance privacy concerns with data usability, emphasizing the importance of metadata for effective discovery.
Data governance and data quality	• Develop standardized approaches and guidelines for metadata management and quality control, alongside technical solutions to tackle data quality issues, fragmentation, and accessibility challenges.

Conclusion and Next Steps

The IDSP convened a virtual roundtable aimed at exploring avenues to propel open data and data sharing practices within the private sector. Through discussions, presentations, and collaborative sessions, participants identified common challenges, shared best practices, and highlighted opportunities in data sharing.

For the next steps, IDSP members will use the feedback and insights from the roundtable discussions to inform the development of a data resource solution by the private sector for the private sector.

Options could include:

- The development of a comprehensive guide or framework, focusing on assessing the value, risk, and feasibility of sharing various types of data, while also addressing legal and ethical implications.
- The creation of a central repository to collate existing resources and to showcase use cases and impact stories of data sharing that resulted in business or societal benefit, or a catalog for discovering, accessing, and collaborating around data.
- The development of a toolkit of available tools and technologies to address privacy and security challenges when making data more accessible.
- The establishment of a community of practice hosted by the IDSP to address topics of interest and exchange best practices.

Appendix

ATTENDEES

Companies represented at the convening were Airbnb, Arup, Bayer, Ford Motor Company, GitHub, HPE, LinkedIn, Microsoft, Northumbrian Water Group, Orange, Ookla, Penguin Randon House, Raven Housing Trust, Roche, Sanofi, Schnieder Electric, Syngenta, Two Sigma, and UK Power Networks.

COMPANY APPROACHES TO SHARING DATA

During the roundtable, companies shared their approaches for how they are making their data more accessible.

- Bayer AG shares clinical trials data, contributing to greater transparency and collaboration in medical research.
- Foundation S – The Sanofi Collective's Open Data for my Child Matters platform aims to facilitate exchange between project teams, and to share scaling and consolidation.
- GitHub shares the Innovation Graph, which analyzes trends in global software development, fostering cross-economy collaboration.
- LinkedIn participates in open data sharing through its Data for Impact program, where aggregated data and insights are shared with public and public benefit organizations.
- Microsoft is committed to making its social impact data as open as possible, with various initiatives outlined in its corporate responsibility efforts.
- Northumbrian Water Group has outlined its approach to open data sharing in its Open Data Strategy, published in June 2023. This strategy serves as a guideline for the company's initiatives in open data.
- Ookla provides data and insights through its Ookla for Good program, supporting organizations and academics in improving internet accessibility for communities.
- Orange adopts a per-project approach for sharing mobile traces processing and analytics data, ensuring quality and usability for purposes beyond network operations.
- Roche demonstrates a commitment to data sharing in clinical trials, promoting transparency and collaboration in healthcare research.
- Syngenta has been actively engaged in open data sharing since 2015. Their approach is detailed in their innovation in agriculture stories, emphasizing the potential of open data in agriculture.

- UK Power Networks centralizes its external dataset publications and implements structure and governance for both open and bilateral data sharing, adhering to regulatory requirements as a utility provider.