ISDM

CENTRE FOR DATA SCIENCE AND SOCIAL IMPACT

The Power of Data for Impact

Part 2 : An overview of data enablers in the social sector

May 2024

Indian School of Development Management, supported by Citi India's CSR efforts, launched the Centre for Data Science and Social Impact (CDSSI) to connect Social Purpose Organisations (SPOs) with funders, academia, tech enablers and other experts, to drive rapid adoption of data and data science by the social sector.

Credits

- **Authors:** Anirudh Ojha, Sreenivasa Madenahally
- **Knowledge and Research Support:** Padma Panchapakesan
- **Reviewers:** Swetha Prakash
- **Dissemination and Publication:** Sowmya Rajaram, Bhavna Panda
- **Design:** Rashi Harsha

Acknowledgements

We are grateful to all the participants and organisations for their immense contribution to the social sector, some of which we have showcased in this report.

Table Of Contents

Introduction: Data For All 1

Support For All 2

Shape It With Policy 5

Open-data Initiatives By Government Of India 6

- Government of India – Comprehensive, multi-sectoral data platforms
- Government of India – Sector-specific data platforms
- State and local government open-data platforms

Helping Hand For The Social Sector 8

- Data-enabling initiatives by non-government entities

Open-data Platforms By Non-government Entities 9

Connecting Across Borders 11

Open-data Platforms By Global Organisations 12

Recommendations For Implementing Organisations 14

Recommendations For Enabling Organisations 16

Conclusion 18

Introduction: Data For All

The effective utilisation of data and data-driven solutions in the social sector is facilitated by a host of organisations, policies, and services. Here are four categories of data-enablers that we will explore in greater detail in this section:

- Organisations with diverse orientations — technology consulting firms, data-science companies and data suppliers, and data-focused funders — are driving the adoption and implementation of data-led solutions within the social sector.
- Efforts by the Government of India, along with various state and local governments, take the form of policies and platforms which play a pivotal role in the use of data.
- Private entities contribute resources that support the implementation of data-driven solutions and practices within the social sector, such as software and services provided at minimal-to-no cost to Social Purpose Organisations (SPOs); assistance from volunteers from the private sector; and open-data platforms.
- Global entities like the United Nations are engaged in a range of initiatives aimed at expediting the advance of data science. Open-data platforms housing datasets from around the world also play a crucial role as facilitators.

Data-enabling organisations in India

- Technology consulting firms
- Cloud service providers
- Specialised data science firms
- Philanthropic and CSR initiatives
- Educational and research institutions
- Government-led organisations
- Volunteer-driven organisations
- Data suppliers

Data-enabling measures from a public policy lens

- National policies aimed at facilitating the use of data by social purpose organisations
- Open-data Platforms
  - Government of India's multi-sectoral and sector-specific platforms
  - State and local government platforms

Data-enabling initiatives by non-government entities

- Software, services, and professional volunteering platforms tailored to social purpose needs
- Open-data platforms by private organisations (free-access and paid)

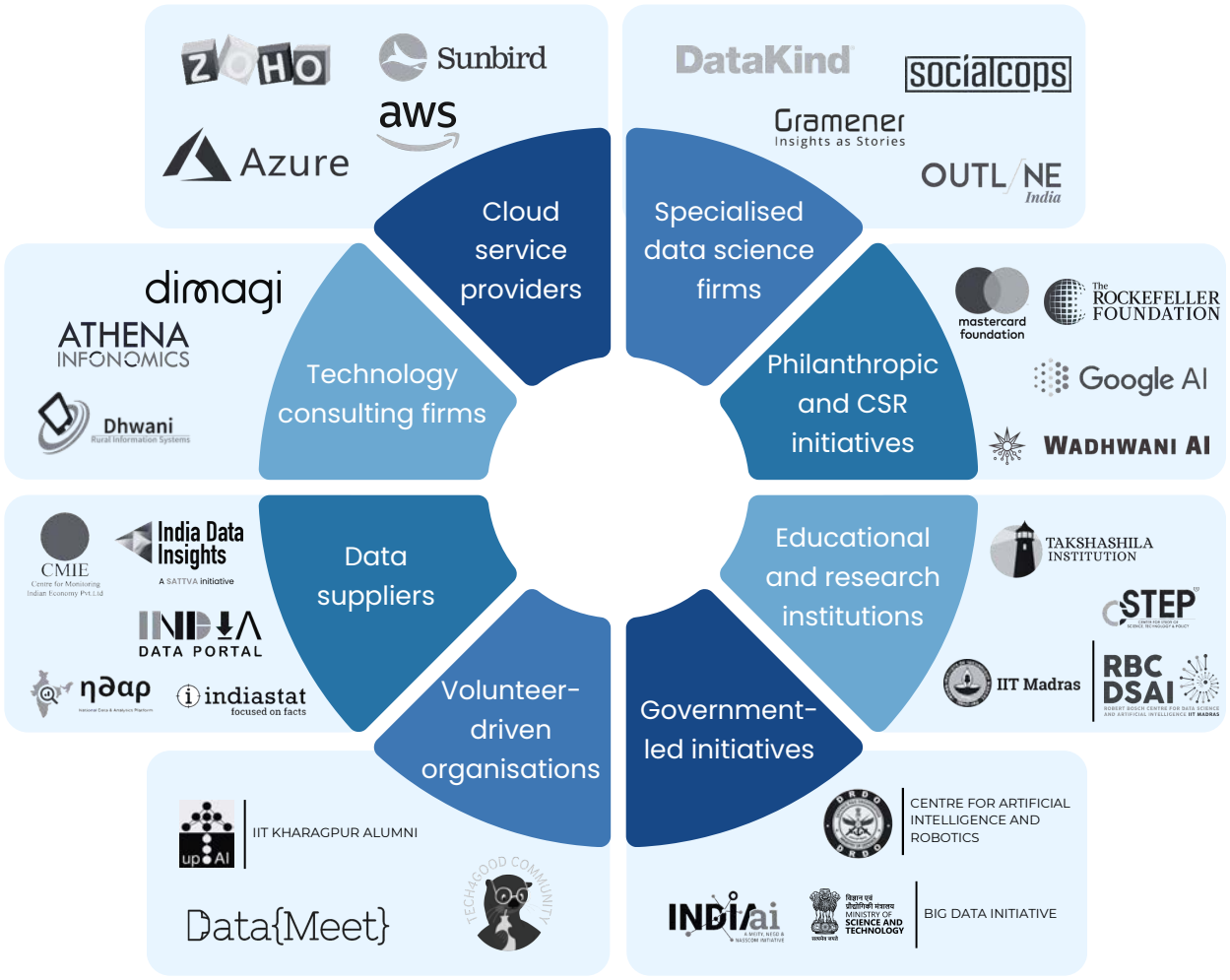
Data-enabling measures from the global social purpose sphere

- Initiatives undertaken by the United Nations for promoting the use of public data
- Global open-data platforms by international organisations

Support For All

Data-enabling organisations play a crucial role in assisting SPOs by providing essential resources, expertise, and support to leverage data effectively. These organisations offer tailored solutions, including technology consulting, specialised data science services, and access to cloud infrastructure, allowing SPOs to harness the power of data for informed decision-making and impactful interventions. Moreover, they facilitate capacity building through training programs and networking opportunities, empowering SPOs to navigate the complexities of data analytics and enhance their operational efficiency, program effectiveness, and social impact.

Some of these entities offer minimal guidance and direction — data suppliers and cloud service providers, for instance — while others, such as tech consulting firms and data science companies, provide more extensive, customised assistance, such as technology consulting firms and data science companies. The following infographic depicts the main categories of data-enabling organisations in India:



Technology consulting firms

Provide tailored data science-powered solutions for specific challenges faced by SPOs.

- Software development
- Implementation
- Managed services
- Human resources
- Technology strategy and planning
- Digital transformations on dashboards to help staff better plan operations.

Cloud service providers

Offer services and infrastructure that enable SPOs to store, manage, and analyse their data efficiently and cost-effectively. They also provide platforms for non-profits to build and deploy applications.

Cloud service providers assist SPOs with cloud-based storage and computing resources through the following:

- Infrastructure as a Service (IaaS)
- Platform as a Service (PaaS)
- Software as a Service (SaaS)

Specialised data science firms

Focus on data science, advanced analytics, and artificial intelligence (AI) in partnering with SPOs to develop and implement tailored data-driven solutions.

- Data strategy and planning
- Data collection and analysis
- Custom AI/ML (machine learning) solutions
- Data audits

Philanthropic and CSR initiatives

Support SPOs in using data science and AI. They provide resources, expertise, and funding to help SPOs adopt and leverage these technologies.

- Funding support for SPOs
- Data science R&D and innovation
- Cost/knowledge/resource-sharing partnerships with non-profit organisations

Educational and research institutions

Conduct research and contribute to advancing knowledge around the use of AI by making their findings accessible to the public

- Joint research projects
- Develop data science solutions that serve as 'proof
- of concept'
- Knowledge dissemination and best practices sharing
- Networks and partnerships

Government-led initiatives

Provide policy guidance, disseminate knowledge derived from research, and work with SPOs to devise data-driven solutions.

- Funding and grants for data-driven projects
- Access to public data
- Policy guidance and support
- Collaboration opportunities and public-private partnerships

Volunteer-driven organisations

Consist of data scientists and other professionals who volunteer their time and skills to help SPOs leverage data.

- Pro-bono data science consulting
- Training and capacity building
- Networking and knowledge sharing

Data suppliers

Public and commercial data suppliers provide SPOs with access to datasets that are used to inform their decision making.

- Access to datasets relevant to the work undertaken by SPOs

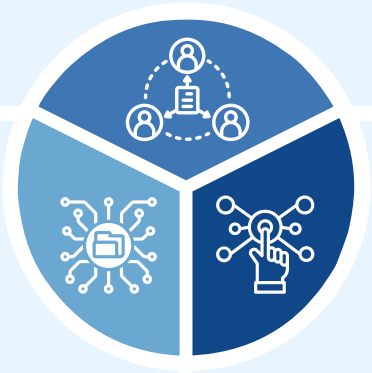
Shape It With Policy

Continuous efforts by the Government of India and various state governments have made data for public use easily available and accessible. The idea is to foster seamless, free, and open access to data across multiple sectors to "...help government, academia, businesses and civil society utilise the available digital data to drive innovation for sustainable development", says a report published by the Ministry of Electronics and Information Technology (MeitY).[1]

Significant improvement in the quality and availability of data offers SPOs the opportunity to include data-driven solutions in their functioning. Some notable policies facilitating this are:

National Data Sharing and Accessibility Policy (NDSAP)

The objective of this policy is to enable access to Government of India-owned data — in both human and machine-readable forms — through a country-wide network in a proactive and periodically-updatable manner, within the framework of various related policies, acts and rules of the central government.



Draft India Data Accessibility & Use Policy 2022

This proposed policy initiative focuses on "transforming India's ability to harness public sector data to facilitate largescale social transformation".[2] It aims to maximise access to — and use of — government data; facilitate the creation of public digital platforms; and protect the privacy and security of citizens. The policy intends to create a balanced and enabling environment for data-driven initiatives, fostering innovation and social impact.

Pillar 6 (Information for All) of the Digital India initiative

The Digital India initiative, launched by the central government, encompasses various pillars to facilitate digital transformation across sectors. Pillar 6, known as 'Information for All', focuses on ensuring the availability and accessibility of data to all citizens. A key element is the 'Open data platform and online hosting of information & documents', which serves as a significant source of information for SPOs.

1. Open Government Data Platform India
2. India Data Accessibility and Use Policy

Open-data Initiatives By Government Of India

The use of high-quality, easily accessible, and freely available data can be transformational for SPOs. Some of the benefits they can derive from open data are as follows:

- Reduce costs: Save time and resources in data collection
- Formulate strategy: Problem identification and definition
- Benchmark: Preliminary assessments and baseline studies
- Showcase: Impact through government-validated data

Government of India – Comprehensive, multi-sectoral data platforms

- Open Government Data Platform
- National Data and Analytics Platform (NDAP)
- Census of India

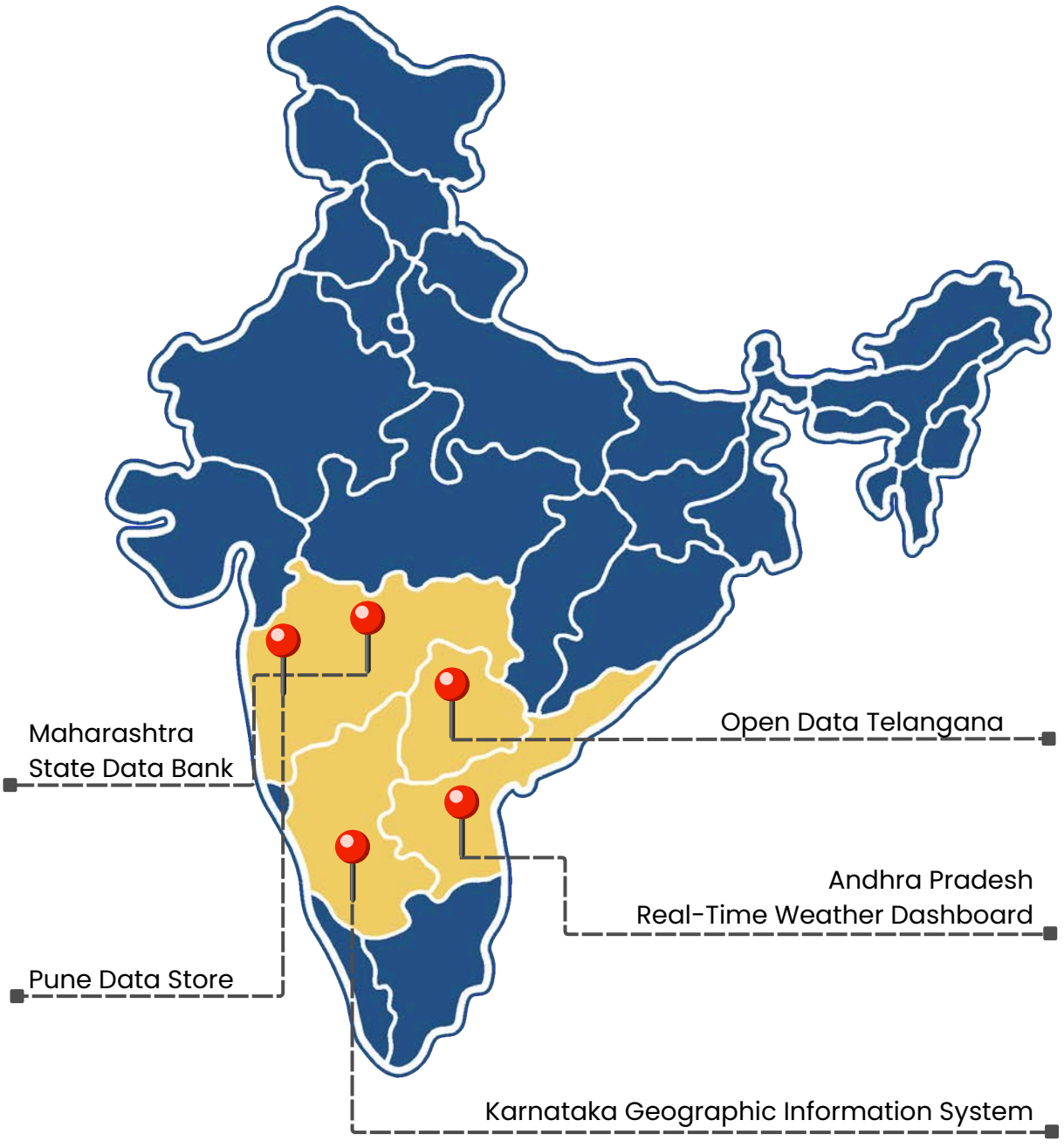
Government of India – Sector-specific data platforms

- Unified District Information System for Education Plus (UDISE+)
- Bhuvan – Indian Geo Platform of ISRO
- National Import-Export Record for Yearly Analysis of Trade (NIRYAT)
- Database on Indian Economy, Reserve Bank of India
- India Urban Observatory
- National Polar Data Center

Note: *Clicking on the platform names will direct you to their respective websites*

State and local government open-data platforms

While most state governments have placed their data platforms on the open platform of the Government of India, a few state governments and municipal corporations have developed their own general and sector-focused open-data platforms. Here are some notable examples:



Note: *Clicking on the platform names will direct you to their respective website*

Helping Hand For The Social Sector

Non-government entities play a vital role in supporting SPOs by offering a plethora of data-enabling initiatives and open-data platforms. These empower SPOs with access to critical information, insights, and the tools necessary for informed decision-making which, in turn, enhances their efficiency and ability to address social challenges.

These initiatives also foster collaboration and knowledge-sharing within the social sector, promoting innovation and collective action towards achieving shared goals. SPOs get access to a wealth of information and datasets, enabling them to gain insights into the communities they serve. Since much of this data is free, it allows them to save on data acquisition fees and reallocate funds to other critical areas.

This section delves into the initiatives introduced by non-government entities towards advancing data-driven solutions and fostering social impact for SPOs.

Data-enabling initiatives by non-government entities

Saathi Re by GiveIndia

This platform provides potential donors with holistic data and insights into non-profits and the impact they create in India. It enables searches based on cause and project across 57,000+ non-profits, operating in more than 35 sectors and across 29 states of India.

DataMeet

This community of data scientists engages in collating disparate government data, converting it into usable formats, and providing previously unavailable data – such as geo-referencing PDFs and image maps of Parliamentary constituencies and village boundaries.

Google for Non-profits

This offers eligible organisations access to Google products that can help solve the challenges non-profits face: finding new donors and volunteers, working more efficiently, and getting supporters to take action. Products such as Google Workspace for Non-profits, Google Ad Grants, YouTube Nonprofit program and Google Maps are offered at no charge.

Online Tech Training Engagement Resource (OTTER) by Tech4Good

This online platform is an aggregator for individuals wanting to contribute to social causes. It brings together software tool makers, investors, NGOs and mentors to help each other create software-aided social organisations that leverage the power of data.

Note: *Clicking on the platform names will direct you to their respective websites*

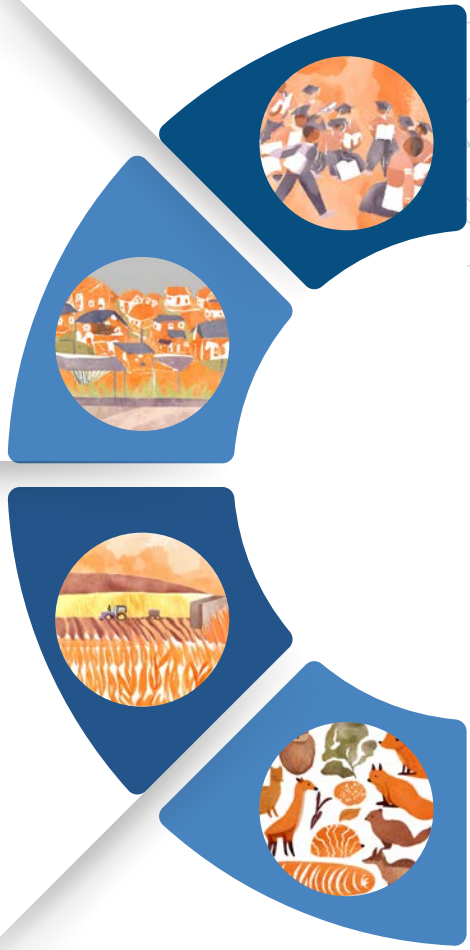
Open-data Platforms By Non-government Entities

ASER by Pratham

The Annual Status of Education Report is a household survey conducted across 616 rural districts in India and covering 6.9 lakh children in the age group of 3 to 16 years to record their schooling status and assess their basic reading and arithmetic skills.

The SHRUG by Dev Data Lab

The Socioeconomic High-resolution Rural-Urban Geographic (SHRUG) platform facilitates data sharing between researchers working in India. This open-access repository currently covers data from 500,000 villages and 8,000 towns.



India Data Portal by ISB

This open-access portal enables users to interact with, and visualise, data and information related to agriculture, rural development and financial inclusion. It contains a data repository with processed and documented public datasets on related themes.

India Biodiversity Portal (IBP)

This portal aims to aggregate information on biodiversity in India and provide free access to this to harness collective knowledge, seek voluntary participation of users and establish a participatory platform for content generation, verification, and usage.

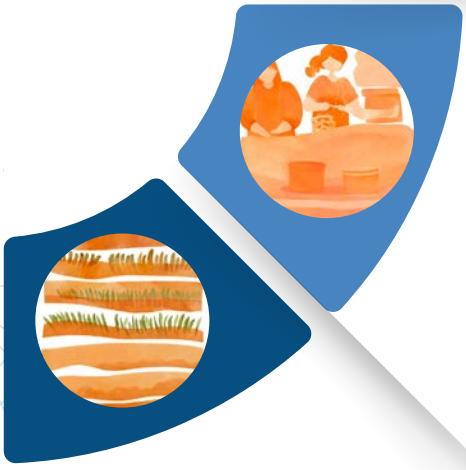
Note: *Clicking on the platform names will direct you to their respective websites*

Centre for Monitoring Indian Economy

The CMIE database provides access to data on various aspects of the Indian economy, including macroeconomic indicators, industry-specific information, employment figures, consumption patterns, and market trends. The data is offered for a subscription.

India Data Insights, Sattva

This open portal provides access to data-driven insights on India's progress on global Sustainable Development Goals (SDGs) and offers readily usable and downloadable charts, reports, and interactive dashboards.



IndiaStat by Datanet

This paid online platform provides comprehensive socioeconomic data on India's sectors, states and regions. The statistical data on IndiaStat pertains to health, agriculture, education, demographics, and social and welfare schemes.

District Level Database by ICRISAT

This database by the International Crops Research Institute for the Semi-Arid Tropics provides information on agriculture and allied services and contains open-access data for 571 districts from 20 states from 1966 to 2015.

Note: *Clicking on the platform names will direct you to their respective websites*

Connecting Across Borders

The integration of data-driven solutions and practices in the Indian social sector is increasingly facilitated by the support of global entities, such as the United Nations and various open-data platforms. These provide vital resources aimed at advancing the adoption of data science methodologies. This offers SPOs the opportunity to leverage data effectively, and empowers them with access to critical information, innovative tools, and collaborative platforms, ultimately enhancing their capacity for informed decision-making and sustainable impact.

This section delves into the spectrum of initiatives and resources offered by global entities, and their impact on the social development landscape in India.

United Nations initiatives for promoting use of public data

The UN often influences the public policies of its member states. It emphasises data-driven solutions and practices to address global/national development challenges. Key aspects of the UN's focus on data include:

Data for Sustainable Development: The UN has placed data at the heart of its 2030 Agenda for Sustainable Development. High-quality, reliable, and timely data is essential to measure progress towards SDGs, identify gaps, and develop evidence-based policies and interventions. The UN encourages member nations to leverage data science to monitor and achieve the SDGs.

Global Pulse: This initiative of the UN Secretary-General promotes awareness of the opportunities Big Data presents for sustainable development and humanitarian action; develops analytic solutions for the UN and its government partners through a network of data science innovation centres, or Pulse Labs; and works to lower barriers to adoption and scaling.

Data For Now: This initiative, tapping into the national statistical systems of member countries, strengthens their capacity to deliver the information needed by policymakers to achieve the 2030 Agenda. It is aimed at supporting member states to leverage innovative sources, technologies, and methods for a streamlined production of data for sustainable development.

Open Data Initiatives: The UN promotes open data as a key enabler of transparency, accountability, and public participation in furthering development. Open-data platforms such as UNData, UN Comtrade, Open SDG Data Hub, and the Humanitarian Data Exchange provide access to a wealth of information, and empower policymakers, researchers, and academics to develop data-driven solutions to global challenges.

Open-data Platforms By Global Organisations

World Bank Open Data

THE WORLD BANK

This portal is one of the largest repositories for global development data. It features a catalogue of time series open data, an online data visualisation tool called DataBank, microdata from surveys, data on world development indicators developed by the World Bank, and more.

WHO Open Data

The World Health Organization publishes health data on this platform, especially pertaining to TB, HIV/AIDS, maternal and perinatal health, mental health, Covid-19 and others. There is also data on popular health indicators such as mortality rates, life expectancy, and number of diarrhoea deaths.

World Health Organization

UNICEF Data Warehouse

unicef for every child

UNICEF publishes child statistics on its open-data platform, including dashboards on topics such as birth registration, child disability, child labour, Covid-19 and even childhood development.

United Nations Office on Drugs and Crime (UNODC) Data

The UN Office on Drugs and Crime conducts research and publishes the data on its platform in the form of research reports, curated statistics, or raw data that can be further analysed. The topics include wildlife crime, homicide, prison statistics, drug use, firearms statistics, and many others.

UNODC

Harvard Dataverse

HARVARD Dataverse

This repository is free for researchers from any discipline, and users can share, archive, cite, access, and explore the data. The Harvard Dataverse is especially useful for SPOs as it holds almost 60,000 datasets in the field of social sciences.

Note: *Clicking on the platform names will direct you to their respective websites*

Demographic and Health Surveys

As the name suggests, this comprises demographic and health data from many countries on nutrition, malaria, HIV, family planning, childhood mortality and much more. There is a good chance SPO will be able to find data from its country in a time-series format from the periodical surveys, helping to make comparisons across time.

OECD Data

This database contains hundreds of statistical time series for the Organisation for Economic Cooperation and Development (OECD) and select non-member countries. It covers broad topics like agriculture, development, education, employment, energy, environment, finance, health, and government.

International Monetary Fund Data

IMF Data is a free and up-to-date resource on all major economic indicators for every country in the world, including debt and capital flows, and a treasure trove of qualitative information about key developments, political and economic, from foreign direct investment (FDI) to conflict and instability.

Statista

This database, in addition to market research data, offers a wealth of demographic information such as age, gender, income levels, education, and geographic distribution. It also offers statistical data on various social and economic indicators, such as poverty rates, employment figures, health indicators, and the environment.

NASA Earth Data

The Earth Science Data Systems (ESDS) program provides full and open access to NASA's collection of Earth science data on the atmosphere, the cryosphere, land, ocean, calibrated radiance, and solar radiance. For SPOs, combining earth science with socio-economic data provides insights for identifying vulnerable communities.

Note: *Clicking on the platform names will direct you to their respective websites*

Recommendations For Implementing Organisations

Focus on building capacity

To enhance capacity and allocate resources, one must ensure that the data and technology requirements align with the mission of the organisation. This involves identifying the specific areas where data and technology best serve the organisation's goals and objectives. Once there is a clear understanding of what needs to be prioritised, allocations for financial and human resources need to be made accordingly. This ensures that resources are directed where they are needed the most, maximising their impact. The organisation can also consider investing in training programs to improve the team's proficiency in using data and technology.

Start with small-scale pilots

Pilot projects serve as invaluable opportunities for SPOs to test the feasibility and effectiveness of data-driven solutions before committing to larger initiatives. By partnering with data-enabling organisations during pilot projects, SPOs can leverage their expertise to design robust data collection methodologies, develop analytical models, and interpret findings effectively.

Leveraging the expertise of data-enabling organisations in their early stages — and in small-scale pilot projects — allows SPOs to minimise risk, learn from the pilot results, and course correct as required. This approach not only helps to refine data-driven solutions, but also builds capacity within the SPO to effectively harness data for future initiatives.

Collaborate with data and tech enablers

SPOs should consider diversifying partnerships by engaging with a wide range of enablers and platforms, including technology consulting firms, data-science companies, and data suppliers, as well as government initiatives for data-enablement. This broad approach lets the organisation explore different perspectives and expertise, enhancing the potential for impactful collaborations. They can also leverage existing platforms and networks to facilitate connections with potential enablers (one such platform is 'Tech to the Rescue', which connects SPOs with tech companies that offer pro bono solutions).

Plan for sustainability

Several key considerations should be made to ensure the organisation's data and technology infrastructure can support its long-term goals. First, aligning the requirements of the data and technology initiatives with the organisation's objectives. This ensures that investments in these areas directly contribute to the organisation's strategic direction and future success. When selecting technological solutions, their scalability must be considered, which involves evaluating whether the chosen solutions can continue to meet the organisation's needs as it grows, so that it can avoid frequent upgrades or replacements. The organisation must ensure that the new data and technology initiatives are compatible with the essential systems and processes already in place. By building upon existing infrastructure, it can minimise disruptions and streamline integration.

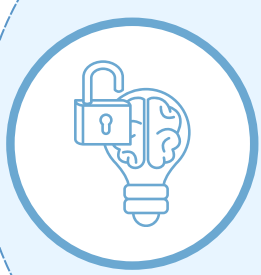
Leverage pro bono support

Leveraging pro bono support from data-enabling organisations is a strategic way for SPOs to enhance their capabilities and maximise impact. They can start by identifying specific areas where data-driven solutions could benefit the organisation, such as program effectiveness, operational efficiency, or impact evaluation. They can also reach out to data-enabling organisations with a track record of supporting social impact projects, clearly articulating the organisation's needs and the potential social impact of the project. Once a partnership is established, they must maintain open communication and collaboration to ensure the effective execution of the project, and capitalise on the expertise and resources provided. By leveraging pro bono support, SPOs can access valuable knowledge and tools to drive meaningful change and achieve their mission more effectively.

Recommendations For Enabling Organisations

Promote pro bono data consulting

Ecosystem enablers can promote the use of data science within the social sector by granting SPOs, particularly non-profit organisations, access to pro bono data consultants who can offer specialised assistance for their projects. This can involve initiatives such as hackathons, mobilising data scientists to tackle specific social challenges, and establishing a dedicated platform for connecting data scientists with social impact endeavour or non-profits seeking data-related support.



Capacity-building to unlock potential

Organisations can offer technical assistance and training programs to SPOs to build their data management and analysis skills. This can include workshops, seminars, and webinars on data collection methodologies, data quality assurance, analysis techniques, and data-driven decision-making. By empowering SPOs with the necessary skills and knowledge, enabling organisations can help them utilise data effectively.

Many SPOs collect a vast amount of data for reporting and compliance purposes. Enabling organisations can help them recognise the untapped potential of existing data, and inspire them to embrace data-driven decision-making.

Assess and enhance data maturity

Organisations can assist SPOs in evaluating their current level of data maturity and identify areas for improvement. They can offer frameworks that measure their data management practices, data quality, analytics capabilities, and organisational culture. They can also provide support in developing a roadmap to enhance the SPOs' data maturity, including recommendations for technology adoption, data governance, skill development, and process optimisation.

Share success stories

Enabling organisations can create platforms for sharing best practices and case studies that highlight the impact of data-driven approaches in the social sector. They can organise conferences and knowledge-sharing sessions where SPOs can learn from each other's experiences and gain insights into successful data initiatives. By sharing knowledge and experiences, enabling organisations can inspire and guide SPOs on their data journey.

Promote data standardisation

To enhance the usability of disparate datasets from various sources, it is critical for SPOs to adhere to consistent metrics, norms, and protocols. By ensuring widespread adoption of standardised measures and procedures for data collection and presentation, they can foster a more cohesive approach to data analysis. This, in turn, sets the stage for eventual data harmonisation, enabling seamless integration and comparison of diverse datasets.

Conclusion

Data and technology enablers play a key role in helping SPOs use data effectively. Enablers could be organisations, individuals, policies, and services, all of which create the conditions for delivering data-driven solutions.

SPOs of all sizes can benefit significantly with help from enablers. Our recommendations [on pages 16 & 17] offer suggestions for how SPOs can leverage these enablers effectively. Most important, implementing organisations should ensure alignment of their data initiatives with their mission, allocate existing resources effectively, and minimise risks by starting small. They should also plan for enhanced capability by leveraging impactful partnerships and pro-bono support. These suggestions are designed to help implementing organisations navigate the complexities of data usage and make the most of available resources.

The report also offers recommendations for enablers to help SPOs better. They can provide much-needed guidance as well as resources that let SPOs maximise the impact of their work. They can support the social sector at large by offering pro-bono consulting, specialised assistance and help in capacity building.

What's next?

In Parts 1 and 2 of the report titled 'The Power of Data for Impact', we provide an overview of the utilisation of data-driven solutions and practices by SPOs. While data and data science have the potential to improve operations and increase the effectiveness of SPOs, their adoption has been limited. This is partly due to the lack of awareness among SPOs of existing gaps in the way they collect, store, manage and use data. For effective utilisation of data, SPOs first need to assess the current health of their data practices and capabilities, using tools such as CDSSI's Data Maturity Assessment (DMA) framework.

The DMA is the next logical step in CDSSI's objective of helping SPOs think of data use in their organisation. Currently, many SPOs are unaware of their data maturity level and how they match up to others in the social sector. To help them identify this, CDSSI aims to implement a Data Maturity Assessment survey for the Indian SPO ecosystem that will:

- Create a deep understanding about the data readiness of SPOs.
- Convey this understanding to SPOs, while also providing solutions and resources that will allow them to plan, prioritise, strengthen and move along this data maturity journey in line with their specific requirements.
- Share learnings and equip funders, ecosystem enablers and implementers to direct their efforts towards addressing data usage gaps (funding, capacity building, consulting).
- Start a conversation with the social purpose ecosystem on data maturity and how to take better-informed, more effective organisational decisions.

ISDM

CENTRE FOR DATA SCIENCE
AND SOCIAL IMPACT