Task Force 05
INCLUSIVE DIGITAL TRANSFORMATION

BUILDING AN INTERNATIONAL INFORMATION ARCHITECTURE: LEVERAGING DATA FOR SUSTAINABLE DEVELOPMENT

Camila Cavalheiro Ribeiro da Silva, Mobility Project Coordinator, Instituto Cordial (Brazil)

Benedikt Gräler, Managing Director, 52°North Spatial Information Research GmbH (Germany)

Luis Fernando Villaça Meyer, Operations Director, Instituto Cordial (Brazil)

Pedro Falha Saraiva, Urban planning Intern, Instituto Cordial (Brazil)

Artur Kim Shum, Innovation designer, Reedu GmbH & Co KG (Germany)

Beatriz de Jesus Gonçalves Francisco, Urban planning Analyst, Instituto Cordial (Brazil)

Doralice Batista das Neves Ramos, Health Analyst, Instituto Cordial (Brazil)

Felix Erdmann, Developer, Reedu GmbH & Co KG (Germany)

Guilherme Pedroso Nascimento Nafalski, Health Project Coordinator, Instituto Cordial (Brazil)

Gustavo de Araújo Barros, Urban planning Intern, Instituto Cordial (Brazil)

João Pedro Silva Araujo, Data Engineer, Instituto Cordial (Brazil)

Rafaela Crossati, Health Analyst, Instituto Cordial (Brazil)

Thomas Bartoschek, CEO, Reedu GmbH & Co. KG (Germany)

TF05

Abstract

The SARS-CoV-2 pandemic, spanning from 2020 to 2023, resulted in over 6.9 million deaths globally, underscoring the necessity of sharing data across borders efficiently to propose timely international solutions. In Brazil, data blackout issues in the Ministry of Health prompted the formation of a Consortium of Media outlets to collect and disseminate disease information. This challenge reflects a broader lack of cohesion in public data collection and availability, evident across various sectors like health and mobility, as indicated by the Covid-19 Transparency Index. Initiatives like the Single Mobility System (SUM) aim to centralize the transport system, enhancing federal involvement and improving data management. Notable practices include DataSUS, a portal aggregating data from the Unified Health System (SUS). The standardization of technologies can streamline data utilization, currently hindered by varying protocols and database structures, impeding interoperability. Embracing digital transformation and platformizing public services as a digital twin of the environment and society can foster inclusion, equity, and sustainable development, reliant on well-structured and standardized data. The G20 must lead a multi-sectoral dialogue, invest in digital capacity building, foster international cooperation and strengthen global data governance to build an international information architecture that promotes innovation and supports human rights. This aligns with the G20's commitments, such as fighting inequalities and promoting sustainable development, strengthening global governance institutions to deal with contemporary challenges.

Keywords: data management; standardization of technologies; digital capacity building; international information architecture.

Diagnosis of Issue

The world faces a vast array of problems that go beyond the individuality of each nation: we share a planet full of crises that affect us all in different ways, such as hunger and poverty, aggravated by another crisis, such as climate crisis. The variability with which these crises affect nations is related to the inequality that still plagues the international scene, the different institutional capacities of each country, the visibility they have of their own local challenges and the possibilities for building solutions in joint international efforts. To reverse this scenario, responsibility must be shared in a democratic way, so that national representatives in multilateral institutions such as the UN, BRICS and the G20 itself can be properly heard and prioritized in the search for solutions.

For this cooperation to be possible, data management must be a natural process and cover the thematic multiverse that encompasses not only the crises we have been through, but also the data access challenge, established by the dichotomy between maintaining each nation legal framework for data protection and privacy, and also in enabling its access. This information can be disclosed to guarantee the exchange of experiences, and recognition of similarities and particularities between countries, essential to understanding the current global crises.

This data must be accessible not only to government decision-makers for effective policy-making and implementation, but also to academia, civil society organizations, journalists and citizens for knowledge production. The private sector also plays a dual role, both as a data provider and receiver, providing information for government interoperability and accessing government data to develop solutions that benefit society.

In data governance, governments at all levels (national, state, and municipal) are major data sources, collecting essential information for public policy and tax collection. They must maintain sovereignty over national data collection and decide what to share, ensuring each legal framework for data protection and privacy.

A dynamic and diverse approach to data collection and access is essential. By ensuring visibility of challenges and supporting evidence-based decision making, as well as monitoring the effectiveness of policies, G20 countries can work towards building a just and sustainable world.

However, we do not live in this reality. Many countries struggle with data issues, including metadata, infrastructure, and operationalizing collection, use, and transparency. Consequently, sectors fall into informational darkness, becoming invisible to local management and hard to compare internationally.

There are several initiatives that seek to solve part of this problem, for example the International Classification of Diseases codes determined by the World Health Organization (WHO), which make it possible to create comparable records and indicators in data systems from different nations. However, there are few initiatives to standardize the structure of data systems that use these classifications in their records on a wide range of topics. Even less so, there are no global standardizations or recommendations that make it easier to structure these data in such a way that they can be related and compared in a large international data system.

A practice that illustrates a relatively integrated system is Brazil's health data system, DATASUS, where there is a centralized department linked to the Ministry of Health that receives information from public facilities in all regions of the country, with the basic aim of distributing resources according to the productivity of each facility. However, obstacles arise in the day-to-day collection of this data:

"inaccuracies in entering data into the system, absence of essential information, lack of individual patient details, impossibility of tracing the patient's trajectory, inadequate grouping of ages and limitation of data to hospitalizations financed by the Unified Health System" (Yano et al. 2021).

Although there is system-level standardization, collection challenges hinder data handling and impact decision-making. Poorly collected data in cities obscures challenges, worsening the crisis. Internationally, the same issue arises, even with initiatives for accessible databases. The Open Geospatial Consortium admits some simple queries are difficult due to massive data. For municipal administrations, information must be effectively available to enable leaders to base proposals on clear evidence for a sustainable society.

A global data architecture speeds up the installation of national data systems, making them comparable across nations. This reduces the invisibility of various challenges and enhances joint solutions, interoperability, and scalable platform services.

Recommendations

A. Survey of Existing Initiatives and Benchmarks

Currently, several initiatives are already aiming to standardize data collection and availability. Some face more or less difficulties in implementation, both in terms of platformization and data collection itself. However, they are already at an advanced stage of implementation.

The G20 should therefore make an effort to carry out a comprehensive survey and diagnosis of these initiatives, at least in the countries of the bloc, in order to ascertain what systems exist and how the knowledge gathered can help in the

creation of a comprehensive Global Data Architecture that builds on similarities and good practices implemented, reducing the risk of technical and operational barriers to integration into a new international system.

There are already a number of initiatives that aim to meet this demand and that present good practices. We can mention national initiatives such as DATASUS, the department of the Brazilian Ministry of Health responsible for organizing the collection of information on hospitalizations and medical procedures throughout the country. Also, the Indian Open Government Data (ODG) and the Health Management Information System (HMIS), both from India[1], and the District Health Information System (DHIS) in South Africa[2], are good examples on a national level.

On the regional scene, we can mention the INSPIRE initiative, coordinated by the European Union and aimed at creating a European Union Spatial Data Infrastructure (SDI) for the purposes of EU environmental policies or activities that may have an impact on the environment, with more than 34 themes, allowing access and sharing between Member States. It was pointed out that interoperability (of data and services) proved to be the essential factor for its full implementation. In general, these and other initiatives seem to be based on the same principles developed by the GOFAIR initiative: Findability, Accessibility, Interoperability, and Reuse of digital assets.

---

[1] The Indian Open Government Data (ODG) platform introduced interaction portals for users to share visuals, studies and reports from the available data, enabling greater engagement with this data culture. In addition, the Health Management Information System (HMIS), also used in India, has an infrastructure based on the specificities of the districts, enabling data to be collected and processed in the various health units and then brought together in a "head quarter".

[2] The District Health Information System`s (DHIS) approach and implementation in South Africa considers local difficulties and the understanding that in LMICS processes of digitalization and platformization of services depend on an organizational and cultural change in the institutions. The initiative has become a national and international benchmark, being adapted and used in other countries.

Even with their challenges, these initiatives are fundamental in the process of disseminating information, and they already have a consolidated structure, both at technical level and in terms of filling out procedures and protocols. They can serve as aggregate knowledge for the production of a global, interoperable data architecture.

B. Structuring a Global Multi-Thematic and International Data Architecture

Identifying topics of public interest common to nations - such as education, health, mobility and the environment, and structuring recommendations for relational database structures, metadata and common technologies, considering standardized indexers, such as the structure of IDs, geolocation and categorical classifications (as in the case of the International Classification of Diseases - ICD), so that all databases can be interoperable and integrated. This structures a unified, multi-thematic and international Global Data Architecture.

This thematic interoperability can take place through Application Programming Interfaces (APIs), which are responsible for allowing the same database to be consulted by different platforms. This, accompanied by the national data infrastructure and authority that will check this data and release it (guaranteeing its processing, anonymity in accordance with national legislation, always regarding what will be made available), tackle the challenge to shared data easily without compromising the national sovereignty. To achieve this, it is imperative that the country establishes a comprehensive and adequate legal framework for data protection and privacy before initiating data sharing abroad.

C. Creation Of an International Data Consortium, Democratically Composed in Order to Equally Manage the System

This whole system, from the drafting of the recommendations to their operation, will need to be managed by the members democratically, fully establishing their rights and duties in terms of ensuring that this information is provided and maintained. Thus, using the same structure proposed by this G20 governance, we recommend that this consortium follows the same rules as the Global Alliance against Hunger and Poverty, to achieve the sustainability, democracy and equity goals set by the G20: National commitments - manages the commitments made by members to adopt effective policies; Financial - has international financial institutions that will provide resources for the implementation of these policies by developing countries; and Technical support - responsible for disseminating and delivering the necessary information and knowledge to member countries.

D. Creation of Working Groups for Specific and Localized Data Processing

With the definition of a data infrastructure, it is recommended that working groups be set up in each member country and responsible for maintaining two pillars in practice: Sovereignty and Individuality. Sovereignty: data doesn't leave the country and isn't distributed outside the government in a raw form. It is always processed, anonymized in accordance with national legislation and made available under the control of the national data authority. Individuality: understanding and comprehending how each scenario behaves with the implementation of this new technological apparatus, as well as the daily collection of information. The intention is for regions that are unable to afford this implementation to be fully assisted in their more specific difficulties, seeking to eliminate the current technological and informational inequality.

The technical and operational support for these cells is represented by the T20 members themselves, who are already working with data, research, and innovations. In this way, legitimized by the governance institution of each state, they can promote a constant flow of work, ensuring the entire functioning of this great system. They also create a permanent binary: each cell made up of both a public and a private Think Tanks, to promote exchanges and avoid conflicts of interest.

In addition, the State can define that one of these cells, linked to its information and data bodies, is in charge of maintaining its DataCenter, since the intention is not for data to be exported from one country to another, which would subject nations to different General Data Protection Laws, as well as raising the issue of sovereignty.

E. Platformization as a Way of Making Information Visible

Finally, it is recommended that this data architecture be connected to a platform capable of delivering this data in a systematic way, taking advantage of technological resources to allow full visualization of the data. Currently, many softwares already offer this functionality, allowing interactive dashboards to be built. This makes it possible to consult everything from local governance, such as city hall, to decision-making by organizations such as the G20.

Building on the concept and ongoing developments of Data Spaces will allow to keep a federated but integrated data infrastructure respecting access and usage rights fostering better data exchange and utilization of available data. It must be stressed that accessibility to digital resources is still a demand in less developed countries, and the International Data Consortium is responsible for providing the necessary inputs so that each member country has the means to implement the technology.

Scenario of Outcomes

The formulation of an International Data Consortium by the G20 represents a significant milestone for the globalized world, offering a series of benefits that would positively impact various aspects of society. Its creation would enable more efficient interaction between member countries, facilitating access to and distribution of relevant data. This would not only promote greater transparency but also allow closer cooperation on issues that require a global approach, such as health, safety, and the environment, without monopolizing the functioning of the structure created and guaranteeing each nation's sovereignty by respecting the legal framework of each country.

In addition, by defining resolutions and recommendations that meet the needs of nations in a multi-scalar way, the consortium would guarantee a comprehensive collection of data and information. By coordinating their respective working groups (cells), made up of Think Tanks that participate in the T20 and are therefore aligned with the values and objectives of this entire project, they make it possible to closely monitor the daily reality of each institution that will collaborate in the collection, crucial for a more accurate understanding of social, economic and environmental dynamics in different parts of the world, allowing for more informed analysis in policy-making and crisis management. The sovereignty of the members over the data would also be preserved, ensuring the trust and continued participation of all those involved, as well as enabling a cultural change in the relationship of institutions with the collection and availability of data, which has already been pointed out by the experience in South Africa as a decisive factor in achieving the

success of a national data system and, extending the logic, to an interoperable international system.

As the main result of the proposal, we can point to the large system created, which includes the data architecture that will allow access and consultation, as well as the set of documents that explain how the information should be collected, separated by theme, thus defining the common indexers and the codes that represent them. National and regional working groups should use these guidance documents to help implement this collection methodology, in order to guarantee the consistency of this data architecture.

Another important benefit would be the feasibility of diagnoses and benchmarks based on the data collected, which would make it possible to monitor the status of member countries in relation to each other more efficiently, comprehensively, and holistically. This would provide a solid basis for defining international cooperation policies and actions while at the same time offering valuable insights for local management, drawing on successful experiences in different parts of the world. In addition to these spaces, this knowledge would also be present in academia, civil society organizations and other environments that not only consult knowledge but also create it, collaborating in the process of monitoring and understanding the different contexts of life in society.

One of the most significant aspects is the creation of a global scenario with greater transparency. By making a broader and more diverse amount of data available, the consortium could highlight countries and regions that currently receive little attention due to the lack of available information. This would promote a democratization of the international debate, allowing previously marginalized

voices to be heard and contribute to the formulation of more inclusive and fair policies.

Establishing a Global Data Architecture and an International Data Consortium by the G20 would mark a major advancement in global governance, enhancing effectiveness, accountability, transparency, cooperation, and inclusivity. This initiative would not only strengthen international relations but also contribute to achieve a more sustainable, just, and equitable world

References

Diário das Leis. "Registro Nacional de Acidentes e Estatísticas de Trânsito (RENAEST)." Accessed May 28, 2024. https://www.diariodasleis.com.br/legislacao/federal/123486-registro-nacional-de-acidentes-e-estatisticas-de-transito-renaest.

GO FAIR. "FAIR Principles." Accessed May 28, 2024. https://www.go-fair.org/fair-principles/.

Government of India. "Open Government Data (OGD) Platform India." Accessed May 23, 2024. https://data.gov.in/.

International Monetary Fund. "e-GDDS Overview." Accessed May 28, 2024. https://dsbb.imf.org/e-gdds/overview.

International Standard Serial Number International Centre. "Standardization." Accessed May 28, 2024. https://www.issn.org/understanding-the-issn/standardization/ .

International Transport Forum. "Transport Data and Statistics." Accessed May 28, 2024. https://www.itf-oecd.org/transport-data-and-statistics/.

Observatório do Conhecimento. "Transparência Covid-19 Brasil." Accessed May 28, 2024. https://transparenciacovid19.ok.org.br/.

Observatório Nacional de Segurança Viária. "Webinar: Análise da Completude dos Dados do Registro Nacional de Acidentes e Estatísticas de Trânsito (RENAEST)." Accessed May 28, 2024. https://www.onsv.org.br/comunicacao/live/webinar-analise-da-completude-dos-dados-do-registro-nacional-de-acidentes-e-estatisticas-de-transito-renaest.

Open Geospatial Consortium. "About OGC." Accessed May 28, 2024. https://www.ogc.org/about-ogc/.

Williamson, L., N. Stoops, and A. Heywood. "Developing a District Health Information System in South Africa: A Social Process or Technical Solution?" *Studies in Health

Technology and Informatics* 84, pt. 1 (2001): 773-777 .https://pubmed.ncbi.nlm.nih.gov/11604842/

Yano, Karen Murakami, Simone Isidoro Prado, Willian Couto, Paola Zucchi, and Maykon Anderson Pires de Novaes. 2021. "Limitações no uso do DATASUS como fonte de dados de pesquisas científicas". Multidisciplinary Health Magazine 2 (4): 27. https://doi.org/10.51161/rems/2780.