Policy Analyses

Building the future of data-driven research and decision-making to control epidemics: a VACCIINE for open data policies.

Authors

Miller Díaz-Valderrama, Ana María-Villadiego, Sebastián León-Giraldo, Natalia Niño-Machado, Swasti Mishra, Cristobal Cuadrado, Catalina González-Uribe, Pasi Penttinen, and Oscar Bernal.

Executive summary

The COVID-19 pandemic has highlighted the critical need for robust open data infrastructures to support data-driven public health decision-making, particularly for controlling infectious diseases. Although researchers recognized the importance of such infrastructures before the pandemic, the global crisis has underscored the urgency of improving data accessibility and quality. Open data policies primarily focus on transparency and accountability but often neglect the practical usability of data for research and public health decision-making.

Open data refers to publicly available information that can be freely used and reused, provided it is transparent, accurate, and respects privacy. Otherwise, open science emphasizes data reuse to enhance research quality and accessibility. It is crucial to differentiate between these concepts and strengthen both approaches to overcome the difficulties of establishing adequate open data infrastructures.

This policy brief is informed by literature reviews, a workshop organized by the Open2Health project and the experience of the Epiverse-TRACE LAC project. It examines the current state of open data for epidemiological surveillance in Latin America and the Caribbean (LAC). It identifies critical barriers and proposes the VACCIINE propositions, a set of guidelines for policymakers to focus on building open data ecosystems for public health. These guidelines emphasize the importance of cultural and infrastructural shifts, including the adoption of standards, improved data architecture, and the role of intermediaries in making data accessible and usable.

Key challenges include distrust, lack of standardization, resource limitations, and the undervaluation of data-related activities. To overcome these barriers, policymakers should prioritize fostering a culture of cooperation over competition, establishing incentives for producing and sharing high-quality data, and acknowledging the contributions of different roles in the data lifecycle. This also includes addressing organizational fears related to data privacy and the potential negative repercussions of transparency.

The brief emphasizes the need to adopt the FREE-FAIRER principles to create a responsive and ethical data ecosystem. These principles focus on making data Findable, Accessible, Interoperable, Reusable, Rapid, and Equitable, ensuring that data serves transparency goals and enhances public health research and crisis response. Policymakers are encouraged to invest in data-sharing infrastructures, incentivize data-related roles, and integrate these standards into public health frameworks to prepare for future health crises and improve public health outcomes.

Description of the problem

Building open data infrastructures to analyse infectious diseases has been a priority since the COVID-19 pandemic. Although its construction attracted researchers' attention years before, it is a necessity that has been widely recognized by academia and decision-makers after facing this global challenge. Still, it remains an outstanding debt (Cetina, 2021). Academia, multilateral organizations, and states ought to join efforts to boost open data infrastructures for data-driven and timely Public Health decision-making, particularly for controlling infectious diseases (D'Agostino et al., 2018).

Data-driven decision-making enables stakeholders to identify and respond to disease trends and patterns. Nevertheless, it is limited by the availability of accurate, diverse, and timely data. Open data infrastructures can help overcome some of those challenges. However, open data policies usually focus more on making open government data available to achieve transparency and accountability and less on making data usable by generating added value through -for instance- research. To enhance epidemic control and analysis, it is essential for stakeholders to differentiate open data from open science and to mobilize efforts to strengthen both because data-driven decision-making takes place between these two ways of openness.

Open data is defined as publicly available information provided in formats that allow for its use and reuse under open licenses and without legal restrictions. This data should be transparent and accurate, and the privacy of sources should be respected (Hrynaszkiewicz, 2013). On a global scale, data openness has been driven by the need for transparency and citizen participation, as well as the potential of technology to improve efficiency and innovation in public management.

On the other hand, open science encompasses a range of initiatives *dedicated to making scientific knowledge openly available and accessible*. It introduces a new framework that incorporates reproducibility, transparency, sharing, and collaboration practices, emerging from the broader accessibility of scientific materials, tools, and methods (UNESCO, 2021). In contrast with open data, open science focuses on the *use* and *reuse* of data more than its *availability*, making *quality* a crucial aspect to consider when releasing information.

In any case, opening is a complex task; different challenges and barriers must be addressed. Distrust, lack of standardization practices, resources, and incentives to make data available and valuable, and the misrecognition of different activities of data work undermine the possibility of going towards building infrastructures for timely research and decision-making on epidemics.

Addressing the question of what policymakers could do to overcome such barriers and help to build such infrastructures, this policy brief aims to provide an overview of the progress and challenges of using data for decision-making in epidemiological surveillance. A literature review on this topic included a revision of lessons from the implementation of open data policies in Latin America and the

Caribbean (LAC). This document is also inspired in the results of a workshop organized in Bogota on June 22 by the Open2Health project, which aims to provide executive training in health governance and crisis resilience, mainly focusing on the role of data in public health decision-making. Additionally, it draws on the experience of the Epiverse-TRACE LAC project, which is oriented toward building an open ecosystem of data tools for the analysis, control, and response to epidemics in LAC.

Open Data for epidemiological surveillance and open science for decision making a (V)(AC)(CI)(INE) formulation

Governments have turned to open data for more than a decade, adopting different open data policies (Attard et al., 2015). Many identified benefits for transparency, accountability, citizen engagement and encouraging innovation may convince governments of these adoptions. However, due to the COVID-19 pandemic, the presence of data gaps and the difficulty of using timely data for decision-making manifested the need for transforming the public access and use of data, especially in health.

Despite *Health* being such an important sector, it remains *data-rich, but information-poor* (OECD, 2022). Thus, governments should direct efforts and investments to build open data ecosystems that could transform data into useful information. Creating open data infrastructures is a step in the right direction; however, it could be futile if the purpose of data reuse does not direct these. The COVID-19 pandemic shows insufficient open government data (LoTempio et al., 2020). There is a need to build open data infrastructure along with open science initiatives, platforms, ecosystems, and guidelines. Collaboration across health authorities within and between countries for this purpose could be strategic. Open data policies should move from mere transparency practices for open government to producing open knowledge of infectious diseases and public health events.

In the region of Latin America and The Caribbean (LAC) by the time of COVID-19 outbreak, different countries had previously started processes of realizing data, regulating, transforming, and producing infrastructures for Open Government. This is evidenced in that, by 2020, at least 1119 commitments made by different Latin American countries members of the OGP (Open Government Partnership) primarily focused on increasing public integrity and transparency (Ramírez-Alujas, 2020). These commitments have led to different initiatives in the public sector to produce regulations, national policies, and infrastructures (such as open data portals). However, they are not focused on research and decision-making.

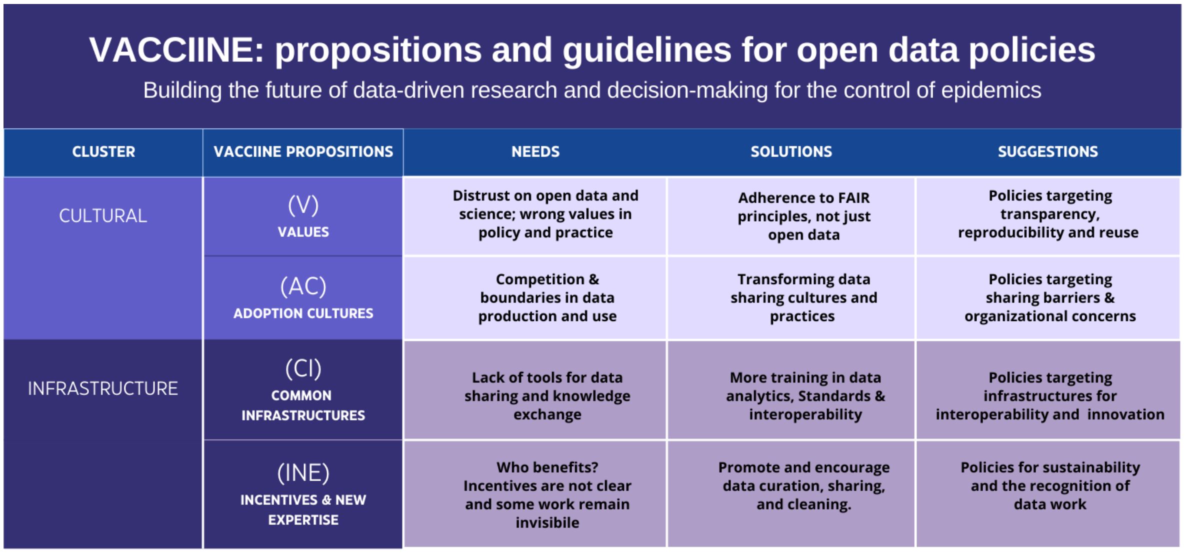
Considering the time, efforts, objectives and scope of regulation, open data policies can be divided into three different waves: the first oriented to the release of data, the second one stimulating its use/reuse, and the third focused on the generation of added value with data (Nugroho et al., 2015). Many of the open data policies in the LAC region have made efforts to release data, and some of them focus on its reuse. However, more efforts need to be made to generate added value.

For the epidemic response and the advance of epidemiological research and decision-making, efforts on funding, infrastructures and initiatives for generating added value are urgent. This urgency is grounded in regional experts' needs and claims who, after facing the COVID-19 pandemic, have gained experience identifying problems and offering solutions on the ground of open data for epidemic control. Needs and claims have been compiled in different documents and papers that provide enough guidance for policymaking and action.

This policy brief supports other policy recommendations for building legal, digital infrastructure, and human capital for data disclosure as a vital need. However, it primarily focuses on identifying fundamental barriers in opening governmental data to generate added value through research and innovation to control infectious diseases. To organize the identified barriers, needs, opportunities, initiatives and suggestions, this document presents the VACCIINE propositions. These topics require pressing attention, and policymakers should concentrate on making data open and valuable for global research and the control of epidemics. However, despite the worldwide character of these efforts, this document mainly concerns the region of LAC and its particularities.

The following synoptic table presents an overlook of these propositions that will be developed in the following pages.

Figure 1-Overview of the VACCIINE propositions synopsis



VACCIINE: propositions and guidelines for open data policies

Building the future of data-driven research and decision-making for the control of epidemics

CLUSTER	VACCIINE PROPOSITIONS	NEEDS	SOLUTIONS	SUGGESTIONS
CULTURAL	(V) VALUES	Distrust on open data and science; wrong values in policy and practice	Adherence to FAIR principles, not just open data	Policies targeting transparency, reproducibility and reuse
	(AC) ADOPTION CULTURES	Competition & boundaries in data production and use	Transforming data sharing cultures and practices	Policies targeting sharing barriers & organizational concerns
INFRASTRUCTURE	(CI) COMMON INFRASTRUCTURES	Lack of tools for data sharing and knowledge exchange	More training in data analytics, Standards & interoperability	Policies targeting infrastructures for interoperability and innovation
	(INE) INCENTIVES & NEW EXPERTISE	Who benefits? Incentives are not clear and some work remain invisibile	Promote and encourage data curation, sharing, and cleaning.	Policies for sustainability and the recognition of data work

The VACCIINE propositions are guidelines to focus policymakers' attention on critical aspects that require urgent action for building open data ecosystems for public health and infectious disease control. These are divided into two clusters of topics, one cultural (purple) and one infrastructural (blue). This synoptic table displays the different propositions regarding identified needs to be approached, the opportunities for change they represent, key initiatives that exemplify possible directions to take and general suggestions for open data policies.

A CULTURAL CLUSTER

Culture influences open data differently and remains one of the most influential factors for adopting open data policies (Haini et al., 2020). Two significant cultural aspects to be transformed are the values that open data policies promote and that need to be developed to make opening data possible and valuable, as well as the organizational cultures in which open data policies and practices are adopted. Policymakers should pay attention to these cultural aspects in formulating open data policies for public health and controlling infectious diseases: the values that orient policies and the organizational cultures that either facilitate or hamper open data initiatives.

Values

Many open data policies in the region of LAC focus on achieving State transparency and open government (Ramírez-Alujas, 2020). Yet, distrust remains a crucial cultural problem to address in public policy and initiatives on open data. It is essential to consider that data disclosure is not the same as transparency, and the latter does not necessarily mean trust. These are often conflated and confusedly interchanged concepts **(Cahlikova & Mabillard, 2020)**. However, it is imperative to call attention to the mechanisms needed to use open data for transparency and to achieve trust.

Although trust building needs to be approached as a complex task (Steedman et al., 2020) that requires actions in the different components of the VACCINE formulation, values are at the core of such an endeavor.

Transparency constitutes the value on which opening data practices rely. However, it could be more explicit. It could be entangled with accountability when public organizations release data to fulfil specific requirements to make State actions available for other organizations and citizens. As a mere requirement of accountable transparency, fear of losing control over data and information and the fear of vulnerability (being seen by others) could influence State organizations to publish raw data in open repositories without considering its quality and usefulness.

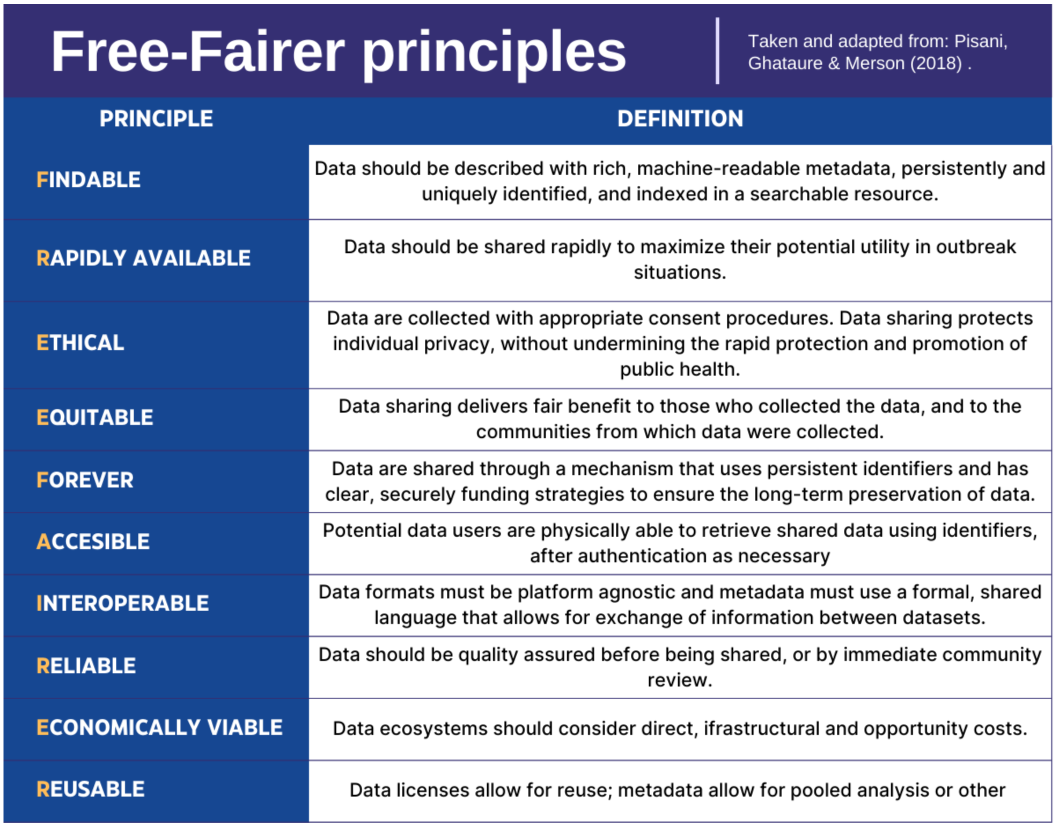
Said otherwise, transparency may need to be adequately followed when releasing data is taken as an accountability practice. As transparency presupposes the capacity for external actors to assess and monitor an organisation's internal workings and performance **(Cahlikova & Mabillard, 2020)**, more is needed. It is mandatory to provide a definition -and to build a practice- of transparency that could include quality and timeliness as essential dimensions.

From the perspective of values, open government initiatives should include open science to not only point at the availability of resources for achieving State transparency but also to power the exploration of data to generate added value. Regarding research, the FAIR principles offer sound guidance for improving the infrastructures that support using open data for research. These principles represent values and guidelines for implementing adequate data management and stewardship (Wilkinson et al., 2016). The FAIR principles include diverse dimensions that should be considered for moving from mere data-releasing practices to the construction of open ecosystems of data.

FAIR is an acronym that gathers the following principles: Findability, Accessibility, Interoperability and Reusability of data. These values focus on building reusable data workflows, highlighting the need for transparency, reproducibility and reusability throughout the data lifecycle. Thus, quality becomes a relevant aspect of open data. What usually happens is that if FAIR principles are disregarded, data availability becomes meaningless, and efforts to facilitate data release may become more expensive. Suppose data quality is replaced by transparency and accountability as values. In that case, organizations can understand data release as the fulfilment of publishing information that is not necessarily good for use in research and decision-making.

As many policies focus on data release to make governmental data available, policymakers must advance further to position the quality at the center of data disclosure efforts. Considering open data for the control of epidemics, this is even more important. To build global health infrastructures for adequate epidemic response, timeliness becomes a significant value in the sense of accurate, high-quality, timely, rapidly findable, accessible, and reusable data. For these reasons, besides the FAIR principles, other aspects should be considered to promote data sharing correctly.

*Figure 2 - The FREE-*FAIRER principles summarized



Free-Fairer principles

Taken and adapted from: Pisani, Ghataure & Merson (2018) .

PRINCIPLE	DEFINITION
FINDABLE	Data should be described with rich, machine-readable metadata, persistently and uniquely identified, and indexed in a searchable resource.
RAPIDLY AVAILABLE	Data should be shared rapidly to maximize their potential utility in outbreak situations.
ETHICAL	Data are collected with appropriate consent procedures. Data sharing protects individual privacy, without undermining the rapid protection and promotion of public health.
EQUITABLE	Data sharing delivers fair benefit to those who collected the data, and to the communities from which data were collected.
FOREVER	Data are shared through a mechanism that uses persistent identifiers and has clear, securely funding strategies to ensure the long-term preservation of data.
ACCESIBLE	Potential data users are physically able to retrieve shared data using identifiers, after authentication as necessary
INTEROPERABLE	Data formats must be platform agnostic and metadata must use a formal, shared language that allows for exchange of information between datasets.
RELIABLE	Data should be quality assured before being shared, or by immediate community review.
ECONOMICALLY VIABLE	Data ecosystems should consider direct, ifrastructural and opportunity costs.
REUSABLE	Data licenses allow for reuse; metadata allow for pooled analysis or other

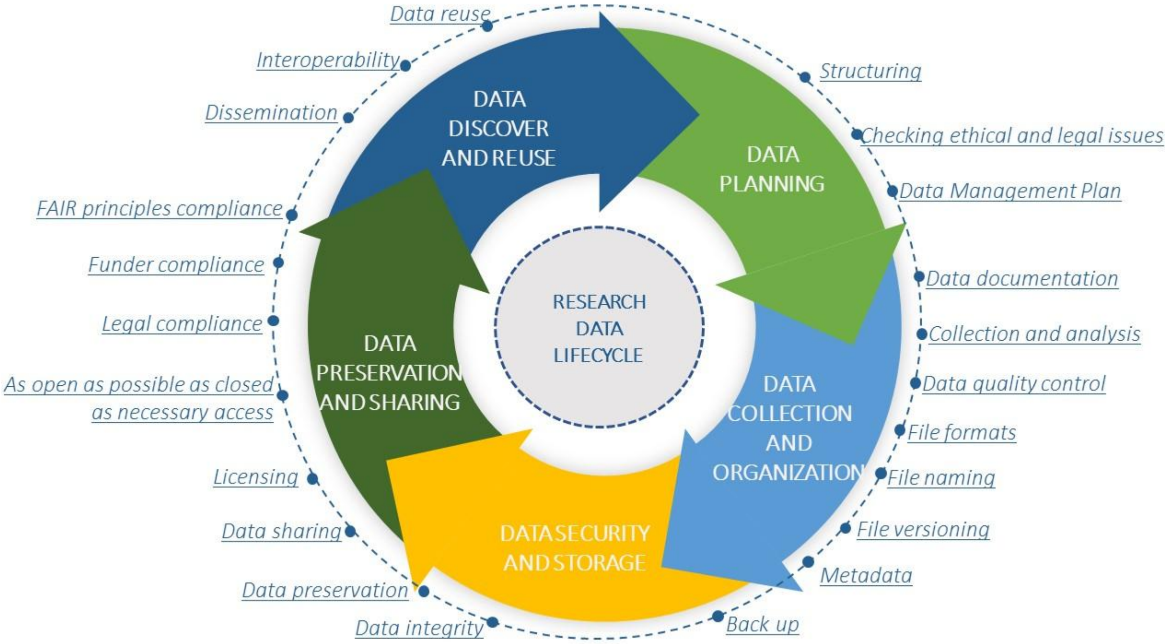
The FREE-FAIRER principles offer a complete set of principles to guide public policy and open data initiatives for emergency preparedness. Besides the FAIR propositions, they highlight the importance of rapid, high-quality, and equitable data as dimensions that few policies consider (Pisani et al., 2018). They also challenge policies to pay attention to the ethical, economic, and sustainability dimensions of opening data.

Policy making following FREE-FAIRER principles could infuse trust in the different actors involved in data sharing for health emergencies. It requires transparency and data reuse as guidelines for action. Thus, it could help close the gap between open government and science efforts.

Adoption Cultures

Even when open data policies are well-oriented regarding values, they can fail once implemented. Besides values, another important cultural dimension to consider is the organizational culture in which closing/disclosing data occurs. Different organizations assume the responsibility of opening data in conflicting ways, for instance, by making essential efforts to publish data in open data portals while maintaining access to metadata limited or paying less attention to the quality of the data they release. Public policies should transform practices and relations with data in the whole data lifecycle.

Figure 3 - Data lifecycle



Direzione Performance, Assicurazione Qualità. Valutazione e Politiche di Open Science. Università degli studi di Milano Via Festa del Perdono 7 20122 - Milano (MI), https://rdm.unimi.it/research-data-lifecycle/

Pressure to release data can sometimes lead to the publishing of low-quality, obscure, or useless information. Four aspects should be considered to limit this risk. One is the relationship between data production, data sharing, and work. It is essential to consider that many valuable public data are the byproduct of activities, such as clinical practice, vaccination strategies, or organizational accountability practices. This implies that physicians, nurses, other professionals, and bureaucrats play a crucial role in the data life cycle, although they often do not recognize themselves as part of it.

On the other hand, public initiatives and policies only sometimes recognize their data work as such. Indifference, then, is the expected response to these exclusions. Thus, there is a need to change the mindset of data producers and policymakers to recognize different activities as data work.

Another critical aspect is the balance between data privacy, organizational secrecy, and the benefits of data sharing. In the context of health and outbreaks, data privacy is paramount, as mishandling sensitive information can lead to violations of citizens’ rights, stigmatization, or loss of trust in public health systems. Policymakers must enforce robust regulations to protect individuals' data while fostering responsible data sharing for public benefit. However, organizations often perceive these regulations as overwhelming, which may delay or deter data disclosure. Organizational discretion, or the requirement for members to keep certain information confidential, further complicates this issue by reinforcing hierarchies between internal and external groups, as well as between governments and citizens. While Open Government initiatives often critique these hierarchies, it is essential to acknowledge that concerns about losing control over sensitive information are legitimate. For example, openness can expose vulnerabilities or errors that may undermine organizational credibility. Platforms like PLISA (Pan American Health Organization's platform for health data) or the IS4H framework demonstrate how digital health transformation can balance privacy with the need for transparency, enabling secure data sharing while maintaining trust and accountability.

Not knowing or committing mistakes are expected conditions of organizational and human work; however, they are highly and costly punished in public administration. Thus, excessive regulation leads to public organizations refraining from data sharing. When regulation is needed, implementing transparent, standardized protocols for managing privacy when releasing information is vital to promote data disclosure practices. On the other hand, as often this organizational fear is accompanied by an understanding of transparency as accountability, policymakers must focus on changing their notions of transparency.

For the case of building open data ecosystems for epidemic control and research, it is necessary to create policies on transparency that privilege replicability instead of mere accountability. Transparency, then, becomes a methodical question: how was data produced? Under which methods and practices was it collected? Which information is needed to make sense of the data? These are questions that the implementation of clear metadata standards could answer. These standards provide researchers and citizens with the necessary knowledge to make data valuable and produce research, innovation and market outcomes (Tenopir et al., 2020).

This movement from accountability to replicability draws attention to another aspect of open data policies: the need for a cultural shift from competition to collaboration. There is strong evidence that the culture of publishing or perishing in academia, which demands researchers to be in a race for publishing results before other competitors, has led to the circulation of distrust among scientific groups who often do not share data before using it. Thus, generating considerable barriers to open science initiatives (Tenopir et al., 2020). This brings attention to the problem of valuing scientific publication as the unique way to publish -or allow publishing- high-quality data. Open data policies should focus on creating incentives for breaking the barriers of ownership in science, especially in topics of great interest, such as population health and epidemic control. Efforts should be made to recognize value in alternative formats and sources for data interchange.

Finally, an issue that open data policies ought to consider is the aversion to *the New* that people in organizations tend to develop when open data threatens to change organizational work. As the construction of open ecosystems of data requires deep transformations, significant opposition is expected. "The argument of "why change what works" is a powerful one regarding every innovation, even more so in the case of reforms that impact the organizational mission and risk upsetting the established routines and work patterns"(Cahlikova & Mabillard, 2020, p. 7). For this reason, each open data policy should be built on historical and contextual grounds, considering the history of trust in public administration in data and science that a particular country has. The history of innovation and State information systems can also inform public policy needs.

Contextual, historical and social knowledge can provide policymakers with enough information to orient their efforts.

INFRASTRUCTURAL CLUSTER

Responding to an outbreak requires sharing readily available information to understand various risks. However, it's crucial also to consider what is not yet known. Not-yet-known information—data in a broader sense—is typically produced in real time and continuously evolves. Infrastructures must be adaptable to extract valuable insights from this emerging data (Heymann, 2020).

In normal conditions, transforming data into information requires a long process involving plenty of heterogeneous actors and activities, including bureaucracy. The importance of open data for the research and control of infectious diseases is the need to shorten this process and to make the times of data production, information, and knowledge sharing closer.

To do so, efforts for data disclosure should be directed at transforming organizational data into digital public goods. Public goods are goods that are both non-excludable and non-rivalrous (Jit et al., 2021). In other words, they should be available and usable for all. Digital public goods encompass ***open-source software,*** open data, open AI models, open standards, and open content. They adhere to privacy regulations and laws, follow best practices, ensure no harm, and support the achievement of the Sustainable Development Goals (Storeng et al., 2021, p. 6). In other words, these public goods build on data to produce social welfare.

Digital public goods differ from commercial digital products. They require a particular ecosystem design, and governmental efforts must be directed to align the complex network of stakeholders, technologies, data sources, and funding organizations embedded in their production (Mikhailov et al., 2024). Mainly, investing in the interconnection and harmonization of infrastructures for data disclosure and sharing for social good is necessary.

Common Infrastructures

It is necessary to build infrastructures to make organizational/governmental action and scientific research move from competition to collaboration practices. This means producing technologies and guidelines that can promote and persuade experts to relate to data as a shared space for innovation and the production of social welfare instead of treating it as a belonging. For this, policymakers should address at least three challenges to transform governmental data into digital public goods: standards, data architecture, and intermediation.

Regarding standards, it is necessary that data releasing could observe precise specifications for making data purposeful (this is an essential condition for data quality). Adherence to the FREE-FAIRER principles and the Digital Public Goods Standard (Digital Public Goods Alliance, 2024) can give open government data enough orientation to avoid releasing useless data and overcome the barriers of distrust. Unclear ownership of data is a common barrier to organizational data release (Beno et al., 2017); considering clear guidelines to allow organizational ownership while avoiding restrictions to access can improve quality data sharing.

On the other hand, the Digital Public Goods Standard highlights the importance of designing and developing open data solutions in compliance with privacy laws and concerns about the responsible use of open data. This is particularly important regarding common fears and tensions arising from using open data and its ethical dimensions. Thus, using these standards can promote the responsible use of data throughout the data life cycle and provide enough orientation to avoid fear.

Concerning data architecture, stakeholders should be granted equal access to data to boost innovation. This requires constructing a robust data-sharing infrastructure that supports equitable access while allowing organizations to maintain governance over their data. Such a system must adhere to ethical standards and prioritize privacy protection, ensuring data sharing fosters transparency and trust without compromising individual rights.

Finally, intermediation. As many open government policies urge public organizations to release data for transparency, the connection between supply and use is usually disregarded. Actors establishing different linkages between these two moments of the data supply chain are open data intermediaries. These vary from open data portals to the development of open software and platforms that give final data users a means to access and use data in a way that fits their needs. They also can be understood as essential stakeholders in the popularization of innovation by connecting actors and encouraging a robust user-provider relationship, facilitating data flow, and targeting new knowledge production (Schalkwyk et al., 2015).

Intermediation is crucial in building open data infrastructures for open science purposes. Data becomes valuable only when it is used or converted into something beneficial. This is possible by producing infrastructures that can process data to provide it to the final user in a form that technical requirements for data use are reduced (Nikiforova & McBride, 2021a), or by guarantying that those with the technical expertise can access metadata and enough methodological information to use datasets effectively. For this, it is crucial that intermediation links such as open data portals could be constructed with a user-centered perspective and regarding protocols for the implementation and evaluation of its usability (Máchová et al., 2018; Nikiforova & McBride, 2021b; Zhang & Xiao, 2020).

Developing and using open software for analyzing and controlling epidemics is vital for real-time data analysis, modelling, and decision-making (Jombart, 2021). Open software enables the integration of diverse data sources, which are crucial for tracking disease spread, identifying hotspots, and allocating resources efficiently. Additionally, the transparency of open-source code allows for peer review, which can improve accuracy and reliability. This openness also facilitates collaboration across borders, enabling global health experts to share insights and develop unified responses to emerging health threats. Therefore, investment in open software for epidemic control should be a priority, accompanied by policies that promote data sharing and collaboration while safeguarding privacy and ethical standards.

Incentives and Expertise

The need for common infrastructures, or technologies that can promote collaboration in data sharing and use is deeply related to the urgency of building incentive schemes for usable data production and release and knowledge sharing. The demand for peer-reviewed journal publications is now a critical barrier to data disclosure. Many scientists do not make their data public because data is jealously guarded either to prevent other scientists from publishing their own or simply because of the fear of getting their findings publicly known before they can publish them. Therefore, the primary way to release knowledge and data about topics of public interest is primarily scientific papers. This long, restrictive and sometimes unequal process impedes different actors from having timely and equal access to data and chances for innovation. Thus, it makes timely responses to expected and pressing issues such as epidemics complex.

The incentives needed should be focused on valuing work that nowadays is not even considered an actual job or essential activity in both science and governmental organizations (Biruk, 2018; Crawford, 2021; Davies & Holmer, 2024; Kingori & Gerrets, 2019). A wide range of activities and people, such as data collection, data cleaning, curation and wrangling, and the different roles entrenched in scientific software production should be recognized and aimed. A meaningful way to do so is by joining efforts with universities, science and technology ministries and organizations to acknowledge quality data production and open release as valuable outcomes. This inter-organizational enrolment is essential to assure the career development of the people involved in different activities in the data lifecycle, not only those who produce new publishable knowledge. Thus, data quality can also be improved.

Creating incentives and recognition for the different expertise in the data life cycle is also crucial for sustainability. Acknowledging the contributions of data managers, analysts, software developers, and others ensures that the entire data production and usage ecosystem is supported. This can be achieved through formal recognition programs, awards, and career advancement opportunities that highlight the importance of these roles. By valuing the diverse skill sets required to manage and utilize data effectively, we can foster a culture of collaboration and continuous improvement. Furthermore, sustained investment in capacity-building initiatives will ensure these skills are nurtured, and best practices are disseminated across institutions and disciplines. In the long term, such an approach will not only enhance data quality and accessibility but also drive innovation and responsiveness to public health challenges and other critical issues.

Conclusion

The COVID-19 pandemic has starkly highlighted a fundamental statement: currently, data quality is far more crucial than quantity. Effective responses to pandemics and epidemics and efficient public health management hinge not merely on data availability but on its accuracy, timeliness, and accessibility, all while upholding stringent privacy and ethical standards. Open data policies transcend transparency when implemented effectively—they become powerful drivers of socially driven innovation and public benefit, particularly in health, where timely, high-quality data can lead to critical, life-saving interventions.

Fostering multilateral collaboration and ensuring equitable data access for all stakeholders is imperative to realize these goals. Adopting the FREE-FAIRER principles—Findable, Accessible, Interoperable, Reusable, Rapid, and Equitable—must be a cornerstone of our approach. This framework amplifies transparency, builds trust, accelerates innovation, and enhances our ability to manage urgent health crises more effectively.

Policymakers must shift their perspective, recognizing data as a vital product that demands a sophisticated infrastructure for effective collaboration. We can establish a robust and resilient public health framework by investing in advanced data-sharing architectures, incentivizing all roles in the data lifecycle, and nurturing a culture of openness and cooperation. Such a framework prepares us for future pandemics and strengthens our capacity to tackle a spectrum of public health challenges.

To build this comprehensive and responsive data ecosystem, policymakers should embrace FREE-FAIRER principles and commit to substantial investments in data-sharing infrastructures. Equitable access and adherence to ethical standards must be ensured. At the same time, incentive structures should be developed to acknowledge and reward contributions across the entire data lifecycle—from collection and cleaning to analysis and software development. This could include formal recognition programs, awards, and career advancement opportunities.

Policymakers must integrate these standards and investments into public health strategies quickly. Implementing these systems can mitigate the urgency of addressing current and future public health crises, ensuring preparedness and resilience. Additionally, ongoing research and policy development should address challenges such as data standardization, privacy protection, and the advancement of real-time data analysis technologies. Prioritizing these actions will enable the creation of a sustainable, resilient data infrastructure, equipping society to confront future pandemics and enhance overall health outcomes and social well-being.

References:

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly*, *32*(4), 399–418. https://doi.org/10.1016/j.giq.2015.07.006

Beno, M., Figl, K., Umbrich, J., & Polleres, A. (2017). Open data hopes and fears: determining the barriers of open data. *2017 Conference for E-Democracy and Open Government (CeDEM)*, 69–81.

Biruk, C. (2018). *Cooking Data: Culture and Politics in an African Research World*. Duke University Press.

Cahlikova, T., & Mabillard, V. (2020). Open Data and Transparency: Opportunities and Challenges in the Swiss Context. *Public Performance & Management Review*, *43*(3), 662–686. https://doi.org/10.1080/15309576.2019.1657914

Cetina, C. (2021). *Policy Brief #28 Integridad y transparencia pública Gobernanza de datos y capacidades estatales para la pospandemia*. https://scioteca.caf.com/handle/123456789/1765

Crawford, K. (2021). *The Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence*. Yale University Press.

D'Agostino, M., Samuel, N. O., Sarol, M. J., de Cosio, F. G., Marti, M., Luo, T., Brooks, I., & Espinal, M. (2018). Open data and public health. In *Revista Panamericana de Salud Publica/Pan American Journal of Public Health* (Vol. 42). Pan American Health Organization. https://doi.org/10.26633/rpsp.2018.66

Davies, S. R., & Holmer, C. (2024). Care, collaboration, and service in academic data work: biocuration as ‘academia otherwise.’ *Information, Communication & Society*, 1–19.

Digital Public Goods Alliance. (2024). *Digital Public Goods Standard*. https://digitalpublicgoods.net/standard/

Haini, S. I., Rahim, A., Megat, N., Zainuddin, M., & Ibrahim, R. (2020). *Factors Influencing the Adoption of Open Government Data in The Public Sector: A Systematic Literature Review*. *10*(2).

Jit, M., Ananthakrishnan, A., McKee, M., Wouters, O. J., Beutels, P., & Teerawattananon, Y. (2021). Multi-country collaboration in responding to global infectious disease threats: lessons for Europe from the COVID-19 pandemic. *The Lancet Regional Health–Europe*, *9*.

Jombart, T. (2021). Why development of outbreak analytics tools should be valued, supported, and funded. The Lancet Infectious Diseases, 21(4), 458-459.

Kingori, P., & Gerrets, R. (2019). The masking and making of fieldworkers and data in postcolonial Global Health research contexts. *Critical Public Health*, *29*(4), 494–507. https://doi.org/10.1080/09581596.2019.1609650

LoTempio, J., Spencer, D., Yarvitz, R., Delot-Vilan, A., Vilain, E., & Delot, E. (2020). We can do better: lessons learned on data sharing in COVID-19 pandemic can inform future outbreak preparedness and response. *Science & Diplomacy*, *9*(2).

Máchová, R., Hub, M., & Lnenicka, M. (2018). Usability evaluation of open data portals. *Aslib Journal of Information Management*, *70*(3), 252–268. https://doi.org/10.1108/AJIM-02-2018-0026

Mikhailov, D., Stewart, U., Hewitt, P., & Porway, J. (2024). *The Data Ecosystem Designer: Designing the Future of Digital Public Goods*. https://data.org/reports/the-data-ecosystem-designer-designing-the-future-of-digital-public-goods/

Nikiforova, A., & McBride, K. (2021a). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, *58*, 101539. https://doi.org/10.1016/J.TELE.2020.101539

Nikiforova, A., & McBride, K. (2021b). Open government data portal usability: A user-centred usability analysis of 41 open government data portals. *Telematics and Informatics*, *58*, 101539. https://doi.org/https://doi.org/10.1016/j.tele.2020.101539

Nugroho, R. P., Zuiderwijk, A., Janssen, M., & de Jong, M. (2015). A comparison of national open data policies: Lessons learned. *Transforming Government: People, Process and Policy*, *9*(3), 286–308.

OECD. (2022). *Investing in health systems to protect society and boost the economy:* https://www.oecd.org/coronavirus/policy-responses/investing-in-health-systems-to-protect-society-and-boost-the-economy-priority-investments-and-order-of-magnitude-cost-estimates-abridged-version-94ba313a/

Pisani, E., Ghataure, A., & Merson, L. (2018). *Data sharing in public health emergencies: A study of current policies, practices and infrastructure supporting the sharing of data to prevent and respond to epidemic and pandemic threats*.

Ramírez-Alujas, Á. (2020). El estado del Estado abierto en América Latina. *Estado Abierto. Revista Sobre El Estado, La Administración y Las Políticas Públicas*, *4*(1), 13–38.

Schalkwyk, F. van, Caňares, M., Chattapadhyay, S., & Andrason, A. (2015). *Open data intermediaries in developing countries*.

Steedman, R., Kennedy, H., & Jones, R. (2020). Complex ecologies of trust in data practices and data-driven systems. *Information Communication and Society*, *23*(6), 817–832. https://doi.org/10.1080/1369118X.2020.1748090

Storeng, K. T., Fukuda-Parr, S., Mahajan, M., & Venkatapuram, S. (2021). Digital Technology and the Political Determinants of Health Inequities: Special Issue Introduction. In *Global Policy* (Vol. 12, Issue S6, pp. 5–11). John Wiley and Sons Inc. https://doi.org/10.1111/1758-5899.13001

Tenopir, C., Allard, S., Baird, L., Borycz, J., Christian, L., Grant, B., Olendorf, R., & Sandusky, R. J. (2020). Data sharing, management, use, and reuse: Practices and perceptions of scientists worldwide. *PLOS ONE*, *15*(3), 1–26. https://doi.org/10.1371/journal.pone.0229003

UNESCO. (2021). *UNESCO Recommendation on Open Science OPEN SCIENCE*.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., … Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data*, *3*(1), 160018. https://doi.org/10.1038/sdata.2016.18

Zhang, H., & Xiao, J. (2020). Quality assessment framework for open government data. *The Electronic Library*, *38*(2), 209–222. https://doi.org/10.1108/EL-06-2019-0145