



FRAMING PAPER

HELLO DATASPHERE

Towards a systems
approach to data
governance

Bertrand de LA CHAPELLE
Lorraine PORCIUNCULA

The Datasphere Initiative is a global network of stakeholders fostering a holistic and innovative approach to data governance. By cultivating dialogue and connecting communities, the Datasphere Initiative connects sectoral silos and people to build a collaboratively governed Datasphere and responsibly unlock the value of data for all.

For more information, visit www.thedatasphere.org or contact info@thedatasphere.org.

REPORT CITATION AND COPYRIGHT

Datasphere Initiative (2022). Hello Datasphere — Towards a systems approach to data governance. Datasphere Initiative. <https://www.thedatasphere.org/>.

The cover illustration has been designed by [Simo Liu](#).

INTRODUCTION

Our data-driven era is fraught with policy challenges. Building a more equitable and inclusive digital society requires a bold change of perspective regarding data governance.

This article presents the concept of the Datasphere as a complex system, arguing that traditional conceptual frameworks and policy instruments are insufficient to govern it, and that innovative mechanisms and institutions are needed. These elements constitute the rationale for the creation of the Datasphere Initiative.

A HUMAN-CREATED GLOBAL DATASPHERE IS SHAPING OUR DIGITAL SOCIETY

The 21st century is — and will increasingly be — the century of data.

Humanity, as well as its systems and machines, collect, process, share and use staggering volumes of digital data, personal and non-personal, public and private.¹ Data now underpins and reflects practically all economic sectors and social activities. Policy-making, at local, national and international levels, also critically relies on data. Data is paramount both to inform individual decisions and to address major global challenges.

Digital data has unique properties, different from traditional goods and services. In particular, its non-rivalrous (yet excludable) nature and unlimited reusability enable increasingly complex value chains, with the potential to create unprecedented social and economic value through sharing.

However, as is often the case with new phenomena, data is misrepresented through incautious analogies (e.g. “data is the new oil”) that inspire ill-adapted policies (e.g. data localization measures).²

The importance of data in our society calls for a more sophisticated conceptual framing. **None of the existing terms dealing with the digital world — such as the internet, cyberspace³, and now the metaverse — properly cover the extremely complex dynamics regarding data happening on multiple levels.** We believe the concept of Datasphere helps approach digital data in a more holistic manner.

The term itself appeared as early as the 1980's, with for instance Douglas Ruskoff describing the Datasphere as “our new natural environment”.⁴ More recently, Grumbach, Bergé and Zeno- Zencovitch (2018)⁵, drawing an analogy with the atmosphere, the hydrosphere and the lithosphere, described the Datasphere as an emergent space, proposing it as the “holistic comprehension of all the ‘information’ existing on earth, originating both in natural and socio-economic systems, which can be captured in digital form, flows through networks, and is stored, processed and transformed by machines.”

Expanding on this latter conception, we propose to define the Datasphere as ***the complex system encompassing all types of data and their dynamic interactions with human groups and norms:***

- **Digital data**, personal and non-personal, private and public, is organized in datasets of diverse sizes and types, although such classifications have blurred, overlapping and moving boundaries. Importantly, the same data can be part of multiple datasets or used in different sectors and the infinite potential for recombination and analysis constantly creates new data or metadata.

¹Data produced worldwide is expected to hit 175 ZettaBytes (ZB) by 2025 (Reinsel, Gantz, Rydning (2018), The Digitization of the World From Edge to Core, <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>). It is estimated that DVDs storing 175 ZB would be long enough to circle Earth 222 times (Marr (2021), How Much Data Is There In the World?, <https://bernardmarr.com/how-much-data-is-there-in-the-world/>.)

²De La Chapelle and Porciuncula (2021), We Need to Talk About Data, Internet & Jurisdiction Policy Network, <https://www.internetjurisdiction.net/uploads/pdfs/We-Need-to-Talk-About-Data-Framing-the-Debate-Around-the-Free-Flow-of-Data-and-Data-Sovereignty-Report-2021.pdf>

³Douzet, F. (2020), Du Cyberespace à la Datasphère. Enjeux stratégiques de la révolution numérique. In Hérodote, Editions La Découverte (pp. 3-14) <https://www.herodote.org/IMG/pdf/edito-2.pdf>

⁴Rushkoff, Douglas (1994), Media Virus!: Hidden Agendas in Popular Culture, Ballantine Books, <https://books.google.com?id=2hpiAAAAMAAJ>

⁵Bergé, J.S., Grumbach, S. and Zeno-Zencovitch, V. (2018), The ‘Datasphere’, Data Flows Beyond Control, and the Challenges for Law and Governance, European Journal of Comparative Law and Governance, Vol. 5, Issue 2, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3185943

- **Individuals and human groups** of all sorts generate, collect, store, process, exchange, make accessible or access, analyze, and use data for various purposes. Distributed across the world, all these actors are interlinked in complex value chains, often with asymmetric power relations.
- A great variety of **norms**, including cultural, legal, and technical ones, set parameters regarding relationships between humans and data, including: high-level principles, international agreements, laws and regulatory frameworks, but also contracts, licenses or terms of service, and even code, standards, and software underpinning technical systems (including that of supporting infrastructures).

This new concept is particularly useful to address the growing dependency of human activities on data. Still, we need to be conscious of the particular characteristics of the Datasphere — and of what these characteristics entail for its governance.

THE DATASPHERE IS A COMPLEX ADAPTIVE SYSTEM WITH EMERGENT DYNAMICS

On an ongoing basis and a global scale, the Datasphere engages billions of actors, whose actions are determined by the norms applicable to them, but also by their personal choices, preferences and interests, as well as the information available to them.

Such a very large number of interconnected agents with the capacity to individually modify their behavior in relation to the environment and the actions of others constitute what the scientific community labels a “complex adaptive system”.⁶ Widely known examples of these types of systems are flocks of birds or schools of fish.⁷ This relatively recent yet powerful field of study now finds applications in an extreme variety of domains, including the environment, social dynamics, evolution, brain activity, or markets, to name only a few.

We postulate here that the Datasphere, as defined above, is a complex adaptive system, exhibiting the well-documented characteristics⁸ of such systems, including: a large number of interconnected agents, non-linear impacts of their actions, positive and negative feedback loops, unintended consequences, structural unpredictability, emergence and path dependencies.

Indeed, a multitude of individuals and organizations guided by legitimate self-interest or malicious intentions can, individually or by aggregation, have a massive impact on the entire Datasphere. Chains of actions and reactions with positive and negative feedback loops produce emergent trajectories for the whole of society, not determined by any superior authority.⁹

Chains of reaction manifest when, for example, governments adopt extraterritorial measures or data localization initiatives for expected short term benefits, incentivizing other governments to adopt retaliatory responses or to replicate such measures, resulting in cumulative negative effects. This can be to the ultimate detriment of the initiator itself. Likewise, influential individuals and very large companies, through the viral propagation of disinformation or the change of a few algorithmic parameters respectively, can impact millions or billions of people, often with unintended non-linear consequences.

Natural systems (e.g. bird flocks or fish schools) are driven by fixed evolutionary rules that animals have no control over. In the **Datasphere, however, public and private actors — or even individuals — have a capacity to intervene at a meta level, to modify the norms applicable to others regarding the collection, processing, access or use of data.** Moreover, evolving social conventions (e.g., political, cultural or even religious) constantly modify actors’ decision-making processes. **This adds new levels of complexity to the behavior of the entire Datasphere, making its emergent trajectory more unpredictable and harder to influence with existing policy tools.**

⁶For a brief introduction to Complex Adaptive Systems, see: <https://www.youtube.com/watch?v=IWhkUne8T68>

⁷See <https://youtu.be/Gjwvsk-6640>

⁸Carmichael, Hadzikadic (2019), The Fundamentals of Complex Adaptive Systems, Complex Adaptive Systems, (pp.1–16) https://www.researchgate.net/publication/333780588_The_Fundamentals_of_Complex_Adaptive_Systems

⁹Schelling, Thomas (1978), Micromotives and Macrobehavior, New York: Norton

TRADITIONAL REGULATORY INSTRUMENTS ARE INSUFFICIENT TO GOVERN THE DATASPHERE

The rapid (and accelerating) technological revolution has major effects on human societies. Yet, no parallel evolution has taken place regarding political and governance tools. This explains the struggle to address the challenges of our data-driven society.

Irrespective of the massive benefits produced by technology and data, legitimate concerns are indeed growing regarding socio-economic development, security and human rights.¹⁰

Replacing the tech-euphoria of the 90's and 00's, a doomsday narrative around digitalization, pervasive in headline news, make it the primary source of deterioration of democratic processes and disruption of the social fabric. Massive surveillance and deepened inequalities within and between countries contribute to this narrative. There is a widespread and deep feeling that the technological system has intractable dynamics of its own and that the overall trajectory of our societies is worrisome.

In the absence of international coordination between governments, let alone consensus among them, governments trying to address these legitimate concerns and reassert control are pushed into adopting unilateral measures through their national legislations. Often under the banner of data sovereignty, numerous uncoordinated measures either aim to extend the territorial reach of their jurisdiction, or impose, directly or indirectly, data localization¹¹ obligations and other licensing provisions.

This legal arms race increases the potential for conflicts between national norms. Furthermore, in the Datasphere, unilateral actions often have consequences opposite to the initial intention. For instance, territorial obligations imposed by a country may limit access to online information and key cloud services or applications by its citizens and enterprises. Likewise, certain measures can create barriers to entry for small and medium enterprises and thus further entrench the position of major players that can bear the corresponding costs. In many regards, data governance issues are "wicked problems" (i.e. problems that generate new challenges when trying to solve them).¹²

In addition, relying on analogies (e.g. data as oil, currency or water) to approach this new complex reality often suggests anachronic regulatory measures, inspired by historical precedents that dramatically overlook the specific properties of data and the Datasphere.

More generally, discussions about data governance take place in multiple sectoral and policy silos, often far from the practitioners or affected agents. Addressing privacy, cybersecurity, national security, content moderation, digital trade, or taxation as fully separate issues may seem efficient in the short term. However, this overlooks major interdependencies. Makeshift efforts to mitigate one concern trigger tensions between different policy objectives and even create additional problems, as the involuntary security implications of strong data protection regimes illustrate.¹³

The hierarchical international system is based on separate nation-states with territorially-defined sovereignty, bound by a principle of non-interference. It emerged in a world with clear frontiers and few cross-border interactions and may remain valid when decision-making only impacts national polities. **But the efficiency and efficacy of this architecture is challenged in global societies connected through online services, where transnational interactions become the new normal, cooperation is required, and emergent dynamics are at stake.**

¹⁰De La Chapelle and Porciuncula (2021), We Need to Talk About Data, Internet & Jurisdiction Policy Network, <https://www.internetjurisdiction.net/uploads/pdfs/We-Need-to-Talk-About-Data-Framing-the-Debate-Around-the-Free-Flow-of-Data-and-Data-Sovereignty-Report-2021.pdf>

¹¹Data-localization measures have more than doubled in four years: in 2017, 35 countries had implemented 67 such barriers, while in 2021, 62 countries had imposed 144 restrictions. (Cory and Dascolly, 2021, How Barriers to Cross-Border Data Flows Are Spreading Globally, What They Cost, and How to Address Them, Information Technology and Innovation Foundation, <https://itif.org/publications/2021/07/19/how-barriers-cross-border-data-flows-are-spreading-globally-what-they-cost>)

¹²For a definition of "wicked problems" and their dynamics, see: Rittel, H.W.J., Webber, M.M. (1973), Dilemmas in a general theory of planning, Policy Sci 4, 155-169

¹³Ashford (2018), GDPR impact on Whois data raising concern, Computer Weekly, <https://www.computerweekly.com/news/252441275/GDPR-impact-on-Whois-data-raising-concern>

INNOVATIVE GOVERNANCE MECHANISMS AND INSTITUTIONS ARE NEEDED

Studies of complex adaptive systems clearly show that they cannot easily, let alone predictably, be handled through rigid external command and control rules.¹⁴ Ensuring that a system remains within acceptable boundaries requires iterative adjustments to the elaborate web of positive and negative feedback loops governing its emergent behavior. This is particularly true in socio-technical systems like the Datasphere, where certain agents (governments but also some private entities) are both internal parts of the system and norm-setters for the behavior of others.

In light of the above, a systems approach to data governance is urgently needed.

A starting point is to identify what the purpose of this governance is. In other terms: **what is the digital society we collectively want to build?**

We believe an overarching common objective should be to maximize the well-being of individuals and societies. This requires broad sharing of data, which creates value, but also to ensure that appropriate resilience mechanisms detect, prevent and remediate misuses. This also entails arbitrating the sometimes conflictual relationships between social and economic value creation and taking into account negative and positive externalities (which are often difficult to quantify).¹⁵ Most importantly, this calls for a fairer distribution of the created value among human groups, within and between countries.

In light of this goal, all actors, irrespective of their nature and purpose, impact and are impacted by the evolution of the Datasphere. This implies, particularly for the most influential of them, a duty to coordinate and cooperate, which does not replace but complements their individual rights to self-determination and agency. This can be encouraged by a proper set of incentives fostering alignment between micro-decisions and the desired macro-behavior.

Yet, a systems approach to data governance should not only rely on expecting changes in attitude, but also on the institutional frameworks that can enable them. As Elinor Ostrom reminded us in her Nobel Prize acceptance speech in 2009: *"A core goal of public policy should be to facilitate the development of institutions that bring out the best in humans"*.

Unfortunately, the current lack of mechanisms to bridge existing silos constitutes an obvious institutional vacuum: a global cross-sector dialogue involving all categories of stakeholders is a priority prerequisite to the transdisciplinary collaboration data governance requires. On that basis, major actors could experiment with dynamic arrangements (e.g. transnational regulatory sandboxes). They could also eventually formalize high-level mutual commitments (e.g. in the spirit of a Framework Convention), that would serve as "focal points" to organize their ulterior independent yet coordinated behavior.¹⁶

More generally, policy-making regarding data needs innovative approaches¹⁷, taking inspiration in particular from agile methods pioneered in software development (iterative and modular steps), but also from systems engineering or biology (activation and repression feedback loops).

Finally, the institutional aspects of a systems approach to data governance should also address the rapidly growing field of technologically-enabled and decentralized bottom-up innovations. Initiatives such as data trusts, fiduciaries, collaborative or decentralized autonomous organizations¹⁸ not only aim to propose solutions to some data-related challenges, but also raise important and novel issues regarding their own governance mechanisms. Ensuring interoperability between a large number of such initiatives may ultimately call for the development of dedicated protocols, like those that enabled the interoperable internet and world wide web, respectively.

¹⁴Miller, John and Page, Scott (2007), *Complex Adaptive Systems*. Princeton University Press

¹⁵MIT Technology Review Insights (2020), *Fair value? Fixing the data economy*, <https://www.technologyreview.com/2020/12/03/1012797/fair-value-fixing-the-data-economy/>

¹⁶In the sense of Schelling. See Schelling, Thomas C. (1960). *The strategy of conflict* (First ed.). Cambridge: Harvard University Press and Wikipedia article at [https://en.wikipedia.org/wiki/Focal_point_\(game_theory\)](https://en.wikipedia.org/wiki/Focal_point_(game_theory)).

¹⁷Japan's Ministry of Economy, Trade and Industry (2021), *Governance Innovation (V.2): a guide to designing and implementing agile governance*, https://www.meti.go.jp/english/press/2021/0730_001.html

¹⁸DAOs: Decentralized Autonomous Organizations, often related to Decentralized Finance (DeFi) and cryptocurrencies

THE DATASPHERE INITIATIVE

Our data-driven era is fraught with policy challenges but no actor or category of actor can solve them alone. Building a more equitable and inclusive digital society requires a bold change of perspective regarding data governance and dedicated efforts towards cooperation.

On this basis, **the Datasphere Initiative¹⁹ is a global network bringing together stakeholders guided by the vision of a collaboratively governed Datasphere. Its mission is to develop agile frameworks to responsibly unlock the value of data for all.**

Through facilitating dialogue, developing evidence-based intelligence, and catalyzing concrete innovations, both technical and normative, the Datasphere Initiative aims to foster the systems approach to data governance that the digital society needs.

CONCLUSION

"The world was so recent that many things lacked names, and in order to indicate them it was necessary to point." This introductory line of *One Hundred Years of Solitude* by Gabriel García Márquez encapsulates the idea that things come into being through words and definitions. Before that, all we can do is either ignore their existence or point at things — and often get misunderstood.

Conceptualizing the Datasphere provides humanity with a shared perspective and helps us become intentional about the digital society we collectively want to build.

Seeing the whole blue planet from space²⁰ made astronauts acutely aware of its frailty and of the limits of the political frontiers we have chosen to organize humanity around. Likewise, the conceptual framework of the Datasphere enables the cognitive shift necessary to develop a systems approach to data governance, re-empower each of us vis-à-vis our common destiny, and proactively assume our collective responsibility to govern this common creation of humankind.

It has not escaped our attention that two of the other major global challenges the world is confronted with, namely climate change and the current health crisis, also need to be — and are in large part — addressed through a systems approach.

It is time for the international institutional architecture to adapt to this changing paradigm. This does not mean replacing existing governance institutions, but rather connecting them across silos and complementing them with agile processes, frameworks, and institutions that foster collaboration and trust.

How we will govern the Datasphere will determine the future of human society in the 21st century and our capacity to deal with global challenges. We can build together a positive vision of what the future of our digital society can be. And make this ambition a reality.

¹⁹The Datasphere Initiative is currently incubated by the Internet & Jurisdiction Policy Network, formally launched in April 2022. More information at www.thedatasphere.org

²⁰For an introduction on the Overview effect see https://en.wikipedia.org/wiki/Overview_effect

thedata sphere.org