

DATA TO THE PEOPLE

How public and open data empower society



EUROPEAN VOICES

EUROPEAN VOICES

European Voices brings together viewpoints from our experts in Europe, working at regional, national and international organization level, on topics that will resonate with public sector leaders across the world.

Evolving legislation and data initiatives. The impact of new technologies on citizen service delivery. Governments' responsibility to enable business and sustainable economic growth in a fair society. Affirming European values in the digital domain. As Europe's digital decade unfolds, diverse points of view across the continent offer valuable insights that can build greater understanding and coherence, while providing a benchmark for the international community.

This series offers comparative perspectives to the global public sector community and a window into the latest thinking shaping local policy, technology choices, and citizen-centric innovation.

ABOUT THIS REPORT

Originally written as "Open data: le temps de changer d'ère" and expanded and complemented through the experience of Capgemini Invent professionals across Europe, this is the first of two opinion pieces. In this first issue, we give European leaders insights into the history, achievements, and immediate future of government open data and public data resources in general. The international community will find inspiration and lessons from the substantial effort that Europe put into developing its policies and technology to make open data not only available but impactful.

The second opinion piece will describe how more mature technology and legal frameworks enable governments to actively lead and support the creation of multiple and interoperable data ecosystems. This happens in the same way they contribute to the country's infrastructure by building a motorway or developing the power network.

More information can be found on [Capgemini Invent's website](#).



CONTENTS

1. Foreword | 04
2. The journey to open data and data ecosystems | 05
3. The impact of open data – making the business case | 06
4. The EU Open Data Directive as a push for interoperability of high-value datasets across Europe | 07
 - Successfully implementing an open data approach – the Lombardy case study | 08
 - New roles and responsibilities | 09
5. The Data Governance Act – a new frontier for European open data | 10
6. Conclusion, References, Authors | 11



FOREWORD

In November 2021, Capgemini participated in the [AI For Good Global Summit](#)¹ organized by the United Nations. The summit aimed to accelerate the achievement of the UN Sustainable Development Goals (SDGs). There is growing awareness of the value of data in supporting a response to the SDGs' range of societal, economic, and environmental challenges. So, it was not by chance that Capgemini chose to focus on the topic of Collaborative Data Ecosystems as a theme at the summit. Whether we are talking about reducing inequality and poverty, ensuring peace and justice, education, or climate change, it is expected that artificial intelligence, algorithms – and technology in general – will play a key role in finding solutions that could make a difference. And guess what? AI has a voracious appetite for data: data to be trained against, data to munch on, data to digest and to transform into insight and opportunities.

But where will all this data come from? How will we ensure that its quality is sufficient and its coverage comprehensive?

The short answer is that a combination of technology, regulation, and public trust will enable the reuse of data traditionally kept close and confidential, and allow us to collaborate and innovate solutions for global and local challenges. New technology is being developed, best practices collected and documented, and infrastructures built to underpin the data ecosystems that are already setting early standards and guidance by players such as Gaia-X², the International Data Spaces Association³, MyData⁴, and Team Data Spaces⁵. Regulation supporting data sharing, particularly in Europe, has deep roots and includes both a mandate to make governments' data available in the open (the objective of the EU's Open Data Directive of 2019 and its predecessors) and the support to enable controlled data ecosystems,

such as in the "data space" model (the upcoming Data Governance Regulation in 2022). What's clear is that we have entered a new era in the use of data.

We intend to offer insight into several aspects of this open data discussion in the context of the creation of European data ecosystems:

- The journey to open data and data ecosystems
- The impact of open data – making the business case
- The EU Open Data Directive as a push for interoperability of high-value datasets across Europe
- The Data Governance Act – a new frontier for European open data

THE JOURNEY TO OPEN DATA AND DATA ECOSYSTEMS

Open data are data that can be freely used, reused, and redistributed by anyone – subject only, at most, to the requirement to attribute its source or to re-share according to conditions that are the same or similar to the original (“share-alike”)⁶. This requirement of unconstrained sharing and reuse means that only rarely can open data describe individuals or include confidential information, but this does not make it any less valuable.

Open data are generally created and offered by public sector bodies as part of their mission and day-to-day operations. Still, nothing stops businesses or individuals from doing the same. For example, most data from EU national statistical institutions are open, as are many national “registers” (such as the reference lists of cities and street names) or geospatial data describing countries (maps, the routes of transport networks, etc.)

Whether you are aware of it or not, open data enables most of the digital services we benefit from every day and is reused, under the hood, by services such as Waze or Apple Maps for our journey planning, or by Google Translate to improve their translation services. Wikipedia is open data: can you imagine a world without it? Open data also makes possible most business-to-business digital services.

In the past, open data was considered primarily a tool for transparency and ensuring the accountability of politicians and public organizations; typically, budget and spending data or voting records were used by engaged citizens to assess the performance of their politicians. Today, open data has a much broader role to play in the common good. Open data propositions are also maturing, with

the focus shifting toward defining success away from volume (typically, the number of datasets available) to “impact” created instead.

To leverage its full potential, open data needs to be embedded in creating new data sharing approaches that consider government, organizations, citizens, technology, and the legal framework altogether as a system, hence the term “data ecosystem”. For policymakers and all other parties involved in the European data ecosystem, there is a need to continually learn from the evolving experience when adopting these new approaches, which will see changes in the standards, governance, policies, and expectations pertaining to open data.

But where did this journey begin? To answer this, we must take a step back in time to the European Union’s Public Sector Information (PSI) Directive of 2003. The Directive regulated the terms of how to make the Member States’ data accessible to their citizens. This also applied to datasets that would normally have had significant commercial value if they were put on the market for a price. To enable access with the minimum possible friction, the Directive established that administrations could only charge a fee to compensate for the marginal cost of publication.

In response to the Directive and its two subsequent iterations, EU countries began to evolve their technology strategies and started offering open data to their citizens on open data portals. In France, for example, the first national portal of this kind was opened in 2011, with the creation of data.gouv.fr. Ten years later, Denmark, France, and Spain have developed significant experience in

this area and rank in the 2020 Open Data Maturity Study⁷ as the top three “trend-setters” in Europe across the four evaluation dimensions of policy, impact, portal technology, and data quality. Nonetheless, progress has not always been as hoped for. We can see that the early momentum has generally slowed. Thus, today the principle of “open by default” is still not fully implemented in most countries. As such, we are not fully exploiting the opportunity to unleash one of the simplest – though most essential – assets for creating economic value, enabling innovation, and making government a trusted partner for its citizens.

Sixteen years since the original Directive, the more recent Open Data Directive of 2019 offers an opportunity for all EU countries to re-energize their open data journeys. It is time to scale the open data reuse cases, such as those for artificial intelligence, to push public authorities forward as role models as providers and re-users of open data in their data ecosystems. With open data offering the potential to inform long-term strategic vision and effective decision-making, we are embarking on the next stage of the journey.

THE IMPACT OF OPEN DATA – MAKING THE BUSINESS CASE

The economic impact of a robust open data policy is significant, and the stakes are high. Our favorite analysis of the impact of open data comes from “The Economic Impact of Open Data” study⁸, developed by Capgemini Invent in the context of Data.Europa.EU, the EU’s program for the promotion of open data.

For **citizens**, transparency was the original building block of trust. Open data from governmental institutions gives citizens a renewed ability to understand public policies, evaluate them and test alternative options. In the Economic Impact of Open Data report, the value of data and the impact it creates is measured in several holistic ways. These include efficiency – with open data saving, for example, 27 million hours in public transport and up to 202,000 lives through faster emergency response; cost savings – with open data cutting, for example, healthcare costs by up to €400,000 due to more rapid first aid by bystanders and up to 20 billion in labor costs due to less time spent in traffic. Emphasizing usage rather than data openness will make it possible to give new meaning to the sharing of public data.

For **governments**, open data allows interaction with an open ecosystem of intelligence around public policies and the building of evidence-based approaches. The Covid-19 pandemic became a demonstrator of this. We could see the importance of systematically collecting data describing the virus, the pathology, and the dynamics of the pandemic to governments and the public. As Capgemini Invent documented in the Open Data Maturity Report of 2020⁹, the need to respond to the emergency led many countries to intensify their efforts. They developed new

initiatives and created means for better communication with citizens, often through data visualizations and dashboards, making the data more easily understandable and insightful.

Curiously, government bodies have been found to be fond re-users of their own open data as much as citizens. Through our relationships with governments’ open data teams, we have evidence of how, over time, the open data portals have become valuable resources for civil servants to discover the data needed.

In **research** circles, whether for business or academic, open data enables and stimulates innovation. It allows the research community to repeat experiments, compare results, refine findings, and update their methodologies accordingly. The research outcome is not just more insightful but also makes the research process and its interests transparent. Research results that are fully documented and supported by the related observations, modeling, and calculations are also enablers for anyone down the line, whether it is further research or application by engineering or to inspire new business.

There is also a good **business** case to be made. The Economic Impact of Open Data report examined nearly 15 benchmark studies and helped establish the median economic impact of open data at 1.19% of the EU’s 2019 GDP. This represented an open data market size of €184.45 billion at the EU level (EU27) in 2019. Crucially, the study forecast a market growth to between €199.51 and €334.20 billion by 2025.



THE EU OPEN DATA DIRECTIVE AS A PUSH FOR INTER-OPERABILITY OF HIGH-VALUE DATASETS ACROSS EUROPE

The EU Open Data Directive of 2019 further supports impact creation by requiring governments to additionally publish the so-called "high-value datasets" (HVDs).

High-value datasets are defined as "documents the reuse of which is associated with important benefits for society, the environment, and the economy, in particular because of their suitability for the creation of value-added services, applications, and new, high-quality and decent jobs, and the number of potential beneficiaries of the value-added services and applications based on those datasets"¹⁰. This is creating a new push, challenging countries to establish infrastructure and processes to make HVDs available free of charge, in machine-readable format, and via APIs (Application Program Interface) where suitable. HVDs will also be, to some degree, consistent across Europe in terms of characteristics, such as granularity, frequency of publications and level of detail, hence enabling essential cross-border interoperability.

The push for HVDs attempts to address how the effective use of open data assets is still hampered by a high degree of fragmentation. Unless there is an explicit intention to make them interoperable, datasets are usually of varying quality, granularity, and scope, using different standards, and therefore difficult to reuse and combine. This means that the potential of high-value data is not being realized.

There are several explanations for this: data is most commonly generated and maintained using IT systems that are themselves heterogeneous, each with its format and specific purpose, and which, over time, failed to value the opportunity of being interoperable with others. The designers and operators of those systems did not foresee multiple uses for the data nor anticipate the prerequisites for sharing and cross-referencing data. In turn, the processes are either expensive or not streamlined, often compromising the rationale or business case for sharing. Inconsistent quality, unnecessarily heterogeneous licensing models, and bad metadata make open data publication, reuse, and discoverability even more complicated. We may say that the battle to make more and better open data available is as old as the battle that our friends in civil

service have been fighting to put their technology in order, within the limitation of budget and of IT talent that is more and more expensive to find.

The challenge for data providers – in and outside government – is twofold. On the one hand, they need to ensure interoperability across all shared semantics and syntax characteristics, a common structure, metadata vocabularies, standards, and cataloging. On the other hand, they need to guarantee sufficient quality, meaning readability, completeness, accuracy, recency and regular updating, coherence, and accessibility. All of this must be done while staying within budget and guaranteeing that mistakes don't compromise confidential data or the privacy of individuals.



Successfully implementing an open data approach: the Lombardy case study

The Lombardy approach is particularly insightful to understand the challenges of open data and data interoperability. This Italian region launched its open data portal as early as 2012. In 2014, the regional government published the first list of datasets recommended for the open data initiative with the regional authorities. However, in 2017, only 18 out of the 1,507 municipalities had provided their datasets.

To improve this situation, the regional authorities launched a funding program to cover the costs incurred by the municipalities. In return, the cities had to commit to setting up an automatic workflow of publishing 10 or 25 datasets from a list of 50 established by the region. The joint funding covered part of the costs of making data available – around €2,000 per municipality, including the human resources needed to prepare the data within each city and the technological solutions required to publish the data. This was a low unit cost, but one that opened the possibility of sharing technological investments and developing solutions that multiple municipalities could reuse. In addition to the funding, the region worked with local stakeholders to establish norms and standards aligned with European guidelines.

The results were encouraging: within one year, the number of municipalities that participated in an open data initiative increased from 18 to 136, with nearly 95 cities publishing the identified priority datasets. Today, there are 5,800 datasets on the portal, which are regularly updated and follow the requisite quality standards for subsequent reuse.

Based on the experience in Lombardy, five factors prove to be essential to promote the publication of open data at the local level: a strong political impetus, a financial incentive to compensate for the costs incurred, establishing norms and standards collaboratively, civil servants with an understanding of data, and sharing technological resources.



Governments must take the next step and scale up their open data use cases

Of course, this begins with tapping into the potential of interoperability by making high-value datasets available across data ecosystems and within open data spaces.

To truly derive value from these ecosystems and open data spaces, there needs to be an investment in talent and new skills. The public sector must present itself as a valid alternative for those professionals whose natural tendency is to turn to the private sector instead, which is more attractive in terms of salaries and is considered closer to the technology field. Training sessions on data and artificial intelligence should be widely implemented for the main public sector stakeholders concerned to familiarize them with the use cases within their own institutions and the need to continue the open data initiatives.

A change in mindset towards data sharing is also essential. For the broader economic and societal good, the public sector needs to be a role model in this respect. To this end, it should adopt an open data mindset and way of working.



New roles and responsibilities

The French Bothorel Report recommends appointing a Chief Officer for Data, Algorithms and Source Codes, positioned in the Prime Minister's Office. In turn, the new role might rely on a network of Ministerial Officers for Data, Algorithms and Source Codes. Successors to the Ministerial Officers for Data, these officers would embody the open data policy at the ministerial level and be equipped with increased human, technical and financial resources. Similarly, in Germany the strategy for federal data and open data addresses the particular roles and training needed to provide sufficient data knowledge and create innovative teams and environments. Competent in both technical and legal matters, professionals in these newly defined roles should be put in charge of both defining the open data policy and negotiating the data sharing terms and conditions with other public authorities and private players.

THE DATA GOVERNANCE ACT – A NEW FRONTIER FOR EUROPEAN OPEN DATA

In November 2020, the European Commission presented the first draft of the Data Governance Act (DGA). The draft regulation aims to facilitate access to public data resources beyond what can be already released as open data. The legislation is expected to be adopted by the European Parliament in early 2022. It will become enforceable in the Member States between 12 and 18 months later, depending on Parliament's final decision. This is not a time for public bodies to rest on the laurels of existing open data programs' achievements. The new legislation will deliver critical changes to the scene we described in the previous pages. Governments need all the time available to make an opportunity out of what, at first, may look like just a compliance matter. As it was for the GDPR, this will be another enabling piece of legislation, clarifying the art of the possible and creating more benefits for citizens and businesses alike.

The Data Governance Act indirectly extends the Open Data Directive requirements by stipulating that public organizations must create the right conditions to share and reuse the data they collect. The scope of the data includes personal data and data protected by intellectual property rights or other confidentiality obligations. An exception is possible for information that could undermine national security.

Similar to the marginal cost that governments can charge for open data, the new datasets falling under the DGA mandate may be monetized to cover the cost of making the data

available, provided that the method of price calculation is published and measures are taken to encourage non-commercial uses.

It might happen that it is not possible to share this data by default. In this case, data providers will have to assist re-users in obtaining consent from individuals or legal entities that may be harmed by sharing their data, as long as such an arrangement does not entail disproportionate costs for the public sector.

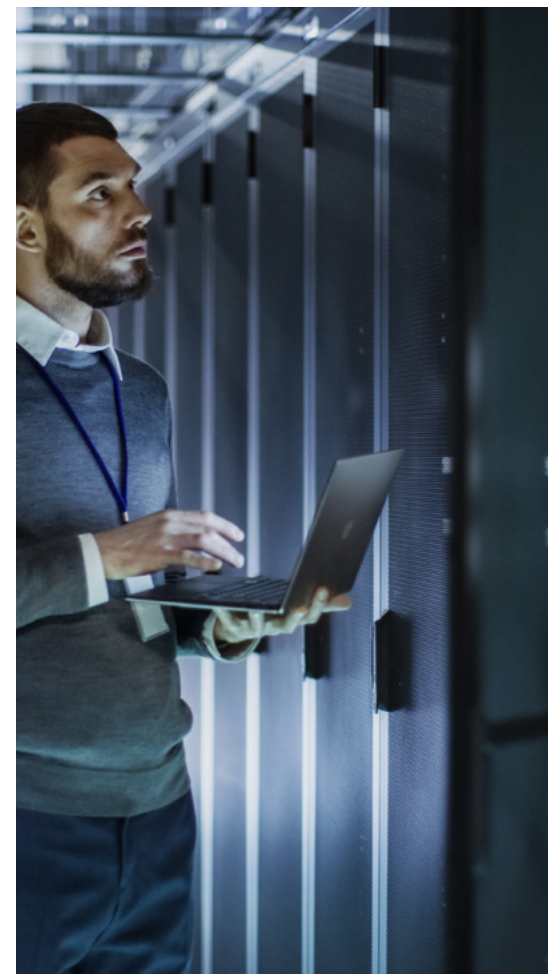
Each Member State will also establish a national authority to ensure compliance with the provisions of the law and support public players in their implementation.

Finally, the regulation establishes a legal regime applicable to all data sharing services. Whether it is the operator of a data hub, a marketplace, cooperatives or trusts, or simply a platform provider for these players, they will be required to notify their activities to the appropriate national authority to perform these services within the EU.

Most importantly, providers of such services will be prohibited from using the data for any purpose other than to improve the data sharing service. The same applies to the metadata collected in the context of these activities. The provider will also have to ensure that the procedure to access their service is fair, transparent, and non-discriminatory to both data owners and data users, including in the context of pricing. In concrete terms, the EU wants to encourage the use of data while preventing the

business value from being leveraged solely by the very few players involved in data sharing.

Providers will have to guarantee "a high level of security for the storage and transmission of data" and "implement the appropriate technical, legal and organizational measures to prevent data transfers or access in cases where it is illegal under EU law", a clear reference to the need to comply with the "Schrems II" ruling of the European Court of Justice.



CONCLUSION

The Open Data Directive reiterated a requirement for the availability and interoperability of government open data that matters most. The upcoming Data Governance Regulation will push the envelope further to data of public interest that can't be opened but, nonetheless, shall not be left underutilized behind closed doors, even when including personal and confidential data or intellectual property. This ambitious target will be achieved by embedding public sector data in the new data ecosystems, whether open or not.

With the Data Governance Act creating renewed momentum, European governments must act without delay to take full advantage of data in transforming public policy and

creating impact at societal, economic, and environmental levels. This means not simply accelerating access to public data; instead, benefitting from the renewed and more mature legal framework, governments must actively lead and support the creation of multiple and interoperable data ecosystems, whether public or private led. The result of all these actions will form what the EU likes to call the single European data space, reminiscent of the single market that the institution's predecessor – the European Economic Community – established in the late 1950s.

The Member States may decide to intervene in the public interest directly by operating the new data ecosystems such as the EU's own Data.Europa.EU

portal for open data, or the French government's Health Data Hub¹¹ for confidential data of patients who are willing to support medical research and entrepreneurship.

Alternatively, governments could enable the private sector to build equivalent infrastructure, which is the most likely model in territories such as the United States. In this latter instance, the imperative is for government to provide guidance as well as regulatory and financial pushes to make the data ecosystems actionable fast. The point of view that follows this will describe how the European Commission is investing serious money in the context of the Digital Europe program to seed the early "data spaces" of this kind.

REFERENCES

¹ <https://www.capgemini.com/fr-fr/wp-content/uploads/sites/2/2021/03/Open-Data-Le-temps-de-changer-dere.pdf>

² <https://www.gaia-x.eu/>

³ <https://internationaldataspaces.org/>

⁴ <https://mydata.org/>

⁵ <https://dataspaces4.eu/>

⁶ <https://opendatahandbook.org/guide/en/what-is-open-data/>

⁷ <https://data.europa.eu/en/dashboard/2020#country-overview>

⁸ <https://data.europa.eu/en/impact-studies/open-data-impact>

⁹ https://data.europa.eu/sites/default/files/edp_landscaping_insight_report_n6_2020.pdf

¹⁰ Directive (EU) 2019/1024 of the European Parliament and of the Council of 20 June 2019 on open data and the re-use of public sector information, Article 2, comma 10.

¹¹ <https://www.health-data-hub.fr/>

AUTHORS

Esther Huyer

Senior Manager
Collaborative Data Ecosystems,
Capgemini Invent Global

Gianfranco Cecconi

Lead for Collaborative
Data Ecosystems,
Capgemini Invent Global

Jean-Guillaume Messmer

Manager
Citizen Services,
Capgemini Invent France

Léo Isaac-Dognin

Director
European Public Sector,
Capgemini Invent France

Pierre Demeulemeester

Vice-President
AI & Data Strategy,
Capgemini Invent France

About Capgemini Invent

As the digital innovation, consulting and transformation brand of the Capgemini Group, Capgemini Invent helps CxOs envision and build what's next for their organizations. Located in more than 30 offices and 25 creative studios around the world, its 7,000+ strong team combines strategy, technology, data science and creative design with deep industry expertise and insights, to develop new digital solutions and business models of the future.

Capgemini Invent is an integral part of Capgemini, a global leader in partnering with companies to transform and manage their business by harnessing the power of technology. The Group is guided everyday by its purpose of unleashing human energy through technology for an inclusive and sustainable future. It is a responsible and diverse organization of 270,000 team members in nearly 50 countries. With its strong 50 year heritage and deep industry expertise, Capgemini is trusted by its clients to address the entire breadth of their business needs, from strategy and design to operations, fueled by the fast evolving and innovative world of cloud, data, AI, connectivity, software, digital engineering and platforms. The Group reported in 2020 global revenues of €16 billion.

Get The Future You Want

Visit us at www.capgemini.com

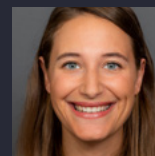
For more details contact:



Valérie Perhirin
Managing Director,
Insight Driven Enterprise - Capgemini
Invent
valerie.perhirin@capgemini.com



Gianfranco Cecconi
Global Lead for Collaborative Data
Ecosystems
gianfranco.cecconi@capgemini.com



Esther Huyer
Senior Manager,
Collaborative Data Ecosystems
esther.huyer@capgemini.com

GET THE FUTURE YOU WANT

The information contained in this document is proprietary. ©2021 Capgemini.
All rights reserved. Rightshore® is a trademark belonging to Capgemini