European Union

Rufus Pollock and Danny Lämmerhirt

Key points

- European Union (EU) member countries have played a leading role in setting the open data agenda with governments pioneering open data policies, licences, and practices in Europe and globally.

- Building on existing public sector information (PSI) policies, significant work has been undertaken related to creating a “digital single market” (DSM); however, open data maturity still varies substantially among countries.

- A number of European governments and the EU have been major funders of open data initiatives, seeking to improve government practices, increase the accessibility and use of government data, and stimulate private sector engagement with open data.

- Basic open government data practices are reasonably embedded in the EU, but with other data-related policy agendas gaining attention (e.g. data protection), and volunteer advocacy in need of new energy, there are threats to the sustainability and progress of open data that will need to be addressed.

Introduction

The history of open data in the European Union (EU) began in the late 1990s and early 2000s. It was at that time that a public sector information (PSI) reuse policy space established the groundwork for more public debate around open data and laid the foundations for key legislative tools in Europe, such as the PSI Directives. Debates at the time revolved around copyright, licensing, pricing, marginal cost recovery, and the political economy of information, as well as the role of citizens and private business in using government data to innovate, especially with respect to government services. Some of the questions at stake included whether government should have a monopoly on data? or should “lean government” provide data to the public to
stimulate economic growth? At the time, this idea of lean government overlapped with other ideas around government reform, such as “government 2.0” or “government as a platform”, a term originally coined by Tim O’Reilly and later adopted by European policy-makers (still present in French open data strategies).³

Although the PSI Directives coexist with other open data policies developed at the national and local level, and individual countries have promoted their own open data principles framed in terms of transparency and accountability, the EU focus on creating a digital single market (DSM) cannot be ignored when seeking to understand the current state of open data in the region. The DSM strategy seeks to enhance the digital economy through cross-country data exchange, envisaging that the reuse of PSI should enable fair market competition. Debates around fair access to PSI and the role of reusable PSI for competitiveness and economic growth overlap to some extent with later open data agendas and policies; however, they also differ because PSI encompasses data and other sources of PSI, such as cultural works. For this reason, this chapter looks in depth at the role of the EU in shaping open data debates.

European countries have been the source of a number of pioneering open data initiatives from the United Kingdom’s (UK) leadership on opening beneficial ownership data to the Government of France’s introduction of policies directing public sector bodies to retain access to data created as a result of public–private partnerships. Key innovations, such as the CKAN data catalogue software, and organisations, such as Open Knowledge International (OKI) and the Open Data Institute (ODI), originated in Europe and have spread across the globe. However, it is also important to recognise that Europe’s policy space developed in parallel with, but was influenced by, legal and policy work in the United States (US).⁴ For instance, decisions in American copyright law influenced the European Database Directive, a key legal feature that influenced open licensing around the world.⁵ Likewise, Gray suggests that the boom of the geospatial data community in the US supported the development of the INSPIRE Directive.⁶

This chapter will start by addressing EU regulations influencing open data before exploring the state of open data infrastructure across the continent and then looking at the ecosystem of actors who have influenced open data development both within the region and around the world.

The European Union as regulator

The Public Sector Information Directive

One key feature of the EU is the directive on the reuse of PSI, also known as the “PSI Directive”. This landmark legislation paved the way for Europe’s open data policies, even though its first version did not refer directly to the term open data.⁷ Also noteworthy are the rules on reuse of the European Commission’s own data.⁸ The Commission states that “the directive on the re-use of PSI provides a common legal framework for a European market for government-held data (public sector information). It is built around two key pillars of the internal market: transparency and fair competition.”⁹

First ratified in 2003, the PSI Directive established a minimum set of rules governing the reuse of PSI. This included PSI at large, covering the “material held by public sector bodies in the
member states, at national, regional and local levels, such as ministries, state agencies, municipalities, as well as organisations funded for the most part by or under the control of public authorities (e.g. meteorological institutes). With recent work on revisions to the directive, some have argued that the directive should be extended further to also cover educational and research institutions. EU member states implement the PSI Directive through different legal tools; some of them ratify specific legislation, others embed the directive in existing access to PSI laws, and some countries use a mix of both approaches. Notably, Belgium developed individual policies for federal and regional levels, while other states issued policies only at the federal level.

Originally only a recommendation, the PSI Directive was revised in 2013 to become an obligation for member states “to make all documents re-usable unless access is restricted or excluded under national rules on access to documents”. This revision should harmonise rules and practices among EU member states. The PSI Directive still leaves flexibility around its implementation, allowing governments to define their own access to document regimes. It is recommended, but not required, that members determine marginal cost recovery regimes for public information, and they are encouraged to use open licences (the effect that this has on open data is discussed further below). In 2013, the directive was also expanded to cover content held by museums, libraries, and archives.

Like every EU directive, the impact of the PSI Directive is assessed at regular intervals. At the time of writing, the PSI Directive is currently undergoing a review that started in 2017. The new revised PSI Directive has not yet been agreed upon, but it is anticipated that several new aspects will be ratified, including:

■ **An expansion in scope.** The new PSI Directive covers data from “public undertakings” (such as government-owned companies), legislative and judiciary government branches, and publicly funded research, although none of these data holders are explicitly obliged to proactively publish data.

■ **Publication of high-value datasets** with a list of specific high-value datasets to be determined in future legal legislation, referred to as “delegated acts”.

■ **Limitation of protection via database rights** such that governments will no longer be able to protect data based on database rights, modifying an important aspect of EU law (see below).

■ **Provisions to cover real-time access to data.**

■ **Curtailing exclusive arrangements** between public and private stakeholders by mandating that governments disclose these arrangements.

However, issues persist. For example, the European Commission has debated what should count as high-value datasets, and some policy-makers considered databases, such as company registers, to be exempt. There are also no public routes to provide input into the designation of high-value datasets. They are likely to be determined by expert groups only. This may raise issues regarding how “high-value” is determined and how diverging values (e.g. economic and social values) will be weighed against each other.
Some voting committees in the European Commission have also sought to change the new PSI Directive by excluding public undertakings as a source of open data, and it is questionable to what extent data from public service providers will be made open in the future. Past shifts in public ownership through public–private partnerships, public sector liberalisation, and the privatisation of key public sector activities, such as the British Royal Mail, have prevented key data from being opened up. This raises questions as to how well existing open data laws can ensure the release of key data for the public’s benefit. France has addressed these issues since 2016 with its “données d’intérêt général” (data of public interest) policy, which contains a clause to enable public sector bodies to demand the provision of data collected from companies that provide public services, such as waste management or transport firms. The ODI has developed a policy guideline on how procurement contracts may include open licensing provisions in order to ensure that key data continues to be available to the public, although there are no numbers on how many contracts contain such clauses.

INSPIRE Directive

The introduction in 2007 of the INSPIRE Directive was another key landmark policy for EU member states. This directive sets minimum conditions for interoperable sharing of spatial data across EU member states and their geospatial communities. The INSPIRE Directive is part of a larger European Interoperability Framework and the e-Government Action Plan that contributes to the DSM.

The Directive is focused on the spatial information infrastructures of the 28 member states of the EU. To enhance interoperability, it requires member states to adopt several measures from data specifications and metadata to data sharing and monitoring requirements. The directive also covers 34 spatial data themes that are relevant for different environmental applications.

The INSPIRE Directive must be fully implemented by all member states by 2021, including requirements to increase sharing services, structure and enhance data, and apply conducive licensing terms. The directive has led to the creation of the INSPIRE Geoportal, a central access point for searching, viewing, and downloading spatial data.

Implementation at the member state level has been mixed. According to a recent status report, national action plans have provided momentum in member states. Many states provided plans for the period 2016–2020 to tackle persistent implementation issues. According to the report, member states need to improve data sharing; strengthen national use cases of spatial data; enhance the management of spatial data, including the identification of spatial datasets; improve access to data; and align spatial data with common data models prescribed by the directive. In particular, spatial data is seen to be underused. The status report suggests that spatial data could be more strongly aligned with other e-government initiatives across Europe.

The Digital Single Market strategy

The DSM is the European Commission’s strategy “to ensure access to online activities for individuals and businesses under conditions of fair competition, consumer and data protection, removing geo-blocking and copyright issues”. More concretely, the strategy seeks to improve
access for consumers and businesses to digital goods and services across Europe, to create the right conditions for digital networks and services, and to grow the digital economy.

Adopted on 6 May 2015, the strategy includes 16 initiatives which were delivered by the Commission by January 2017. The evolution and progress of the DSM strategy can be assessed according to how these initiatives were developed. Achievements under the strategy include the ratification of the General Data Protection Regulations (GDPR). Other legislative frameworks introduced in this chapter also contribute to the DSM. For instance, the PSI Directive creates a common data space in the EU and regulates common access conditions for PSI. The copyright reform (ongoing at the time of writing) may also be a key anchor point for the DSM.

A mid-term review of the strategy in 2017 stressed that Digital Public Services were key to further advancing the strategy. While the focus is on e-government and open data is not mentioned specifically, open government data programmes are likely to play an important role given their close links to other progressive ideas, such as “Government 2.0” and “government as a platform”.

The role of the EU as a funder

The EU plays an important role in funding research and innovation programmes through its Horizon 2020 financial programme. The European Commission publishes the names of the projects that receive funding through the CORDIS database. A query of the CORDIS database, using open data to understand the nature of EU funding of open data research under Horizon 2020, produced a result of 63 projects. Based on query results, € 223 million has been spent on the following topics and applications:

- **Mobility**: Open data protocols to be used for smart mobility.
- **Energy**: Heating and energy management with open data as one enabling component.
- **Pollution**: Publication and use of open air pollution data.
- **Agriculture**: Open data input into agricultural services.
- **Research**: Open science and citizen science, using open data as one component to exchange information across researchers (e.g. one project used open data sources to map political concerns of citizens).
- **Utilities**: Open data for water management.
- **Geo-data**: Open geospatial and earth observation data.
- **Health**: Gathering and publishing open health data and open data on clinical trials.
- **Business models**: Using open data as an enabler for business.
- **Digital humanities**: Digitising text and making it available and processable as open data.
- **Housing**: Developing an infrastructure for open and closed housing data.
- **Innovation**: Open data for public service innovation.
- **Open data**: Capacity building around open data for researchers, open data for smarter decision-making, and research on open data’s contribution to the economy.
For many of these projects, only a small share of the funding focused directly on open data per se. For most projects identified, open data plays an enabling role or is a planned project outcome. This suggests, however, that a substantial number of European projects can, or might, put open data and existing open data infrastructures and frameworks to use.

**Implementing PSI in national open data policies**

Since 2010, open data has become a topic of national debates with member states commissioning studies and white papers, issuing decrees and executive orders, developing national open data strategies and policies, and ratifying laws to act upon new commitments to open up data. The latest report on Open Data Maturity in Europe\(^23\) states that governments have significantly invested in developing open data policies with 27 out of 28 EU member states now having a national open data policy. Twenty of these countries have stated that their policies are more ambitious than the PSI Directive by going beyond the passive availability of data on request by proactively publishing data and making it reusable, as well as by supporting training, awareness raising, and engagement activities.\(^24\) Motivations to develop open data policies range from enabling citizens to exert more power over the state and hold wasteful government to account to creating opportunities for small and medium-sized enterprises (SMEs), enhancing the efficiency of government, allowing for better access to government information, and making more effective use of new technologies in the public sector. This is distinct from the PSI Directive’s focus on fair conditions to access PSI and the creation of an EU-internal digital market. Yet, challenges remain. The PSI Directive has to be embedded in federal and regional legislation, requiring multiple legislative documents to open up data. A particularly complex case is the UK, where, for example, the complex administrative devolution requires the development of a unique open data policy for Northern Ireland’s administrative bodies and parliament.\(^25\) Governments have started responding to this complexity by developing harmonisation agencies (e.g. Belgium’s l’Agence pour la Simplification Administrative\(^26\)), and several cities have promoted coordination with national data portals.\(^27\)

---

**The state of open data in the United Kingdom**

The UK played a key role in promoting open data by launching the data.gov.uk portal in February 2010, listing over 42,000 datasets to date. In 2010, Conservative Party leader David Cameron made open data a part of his electoral campaign and pledged that his government would publish an “unprecedented amount of government data”. As Prime Minister, he spearheaded initiatives such as the publication of the world’s first open beneficial ownership register in 2016.\(^28\)

While the UK has continuously performed well in international open data rankings, challenges still remain. Open data lost most of its momentum in late 2015 as government attention turned to the Brexit referendum and later to Brexit negotiations. Many open data advisory bodies ceased to exist or merged with others. For example, the Public Sector Transparency Board became part of the Data Steering Group in November 2015, and the Open Data User Group discontinued its activities entirely in 2015.\(^29\) There have
also been political attempts to limit the Freedom of Information Act (FOIA) based on the argument that opening up government data would be an adequate substitute. There are still issues around publishing land ownership information across all regions, and some valuable datasets have been transferred out of government ownership avoiding publication, such as the Postal Address File that was sold off during the privatisation of the Royal Mail.

Open licensing

Within existing legal and policy frameworks, most European countries publish at least some key datasets under open licences. However, Europe’s open licensing landscape is fragmented, with several bespoke licences adopted by governments. Critics argue that this “licence proliferation” can increase legal uncertainty and complexity. It may also cause incompatibilities with other open licences and create data use silos. The result would be an incompatible licensing system based on many individual contractual terms that may jeopardise the efforts of the European DSM strategy, prevent the free flow of public sector information, and impede the growth of data economies.

No systematic study has explored the exact extent of this problem, but there are some indications of emerging issues. The European Open Data Portal indicates that there could be up to 90 different licences currently used by national, regional, or municipal governments (although the real number may be lower when cases of the same license under different titles are taken into account). Another report commissioned by the EU suggests that small changes and variations between European bespoke open government licences can cause them to be incompatible.

Database rights: A European legal phenomenon

European legislation has had significant impact on the development of open licences in Europe and worldwide. This is due to a key legislative text called the EU Database Directive. It is a “database right” which protects investments made to build a database. This stands in stark contrast to the situation in regions like North America, where database records or the database structure needs to include works of creative authorship to be protected by copyright. Between 2009 and 2010, the Open Data Commons created global licences to address database rights, including the Open Database Licence 1.0 (ODbL) and Open Data Commons Attribution Licence 1.0 (ODC-BY). Creative Commons updated its Creative Commons Attribution (CC-BY) and Creative Commons Attribution Share-Alike (CC-BY-SA) licences to the worldwide version 4.0 to explicitly address database rights globally.

In order to govern, control, and harmonise the adoption of open licences, the EU and its member states have developed several approaches using policy, as well as governance and educational frameworks. At the EU level, the Legal Aspects of Public Sector Information (LAPSI) project, co-financed by the EU, is noteworthy, containing recommendations around the reuse of public
sector data in Europe. Tools such as the European Data Portal’s Licensing Assistant also help governments to identify compatible licences.

It should be mentioned that the PSI Directive does not require the use of standard open licences, such as Creative Commons licences, which would help further harmonise licensing regimes. National licence regimes are governed by different policy tools (laws, policies, decrees, and executive orders) and other soft policy mechanisms (see box on National governance below). However, some national governments, including Germany and Italy, have updated their bespoke licences to ensure compatibility with the Open Definition and standard open licences.

Some public sector bodies may have business models that make them opt out of using open licences. This is supported by the PSI Directive which allows public sector bodies to charge fees for data reuse at higher than marginal costs (although fees must be limited to the costs of reproduction, provision, and dissemination). Examples of datasets available for a fee include the company register in the Netherlands or land registration data in the UK. In other cases, opening up data is controversially discussed under competition law. One such example comes from Germany concerning the release of open weather data under the “Verordnung zur Festlegung der Nutzungsbestimmungen für die Bereitstellung von Geodaten des Bundes” (GeoNutzV). The case suggests the potential emergence of country-specific data markets, with business associations, sometimes successfully, pushing against open data in order to maintain information oligopolies.

National governance of open licences: The cases of Belgium and the United Kingdom

To harmonise the open licenses used, Belgium has developed open data policies at the federal and regional levels. As part of this work, the government has appointed a committee to support agencies in making better licensing choices through a “comply-or-explain” approach. Belgian open data policies define standard open licences (CC0 and CC-BY 4.0) that are recommended for use by agencies. If an agency still wishes to use a bespoke licence, it must send a justification to the committee, which will review the justification to understand whether a bespoke licence is necessary or not.

The UK has developed several mechanisms to ensure a widespread uptake of its Open Government Licence (OGL). The British government initially faced an issue with its OGL since Crown Copyright only applies to government agencies which have Crown status. This means that public sector bodies without Crown status were not able to use the OGL. Open Government Licences 2.0 and 3.0 address this issue by being applicable to rights beyond Crown Copyright. In order to help government agencies, the British government has also developed a licensing framework, providing educational and training resources for government agencies to help ensure harmonisation of licensing practices.
The unfolding impact of the General Data Protection Regulation

In May 2017, the EU ratified the GDPR. This is a broad data policy spanning topics from consent and privacy by design through to an individual’s rights over information about them, the lawfulness of personal data processing, and the rights and responsibilities of data controllers and data processors.

The advent of the GDPR has sparked several discussions about its potential impact on open data. The GDPR requires any processor to have a legal basis to process personal data, primarily, but not only, by gathering consent. This has prompted questions about the reusability of open data containing personal data, such as company registers or public contracting data that may include names of company directors or the disaggregated salary figures of civil servants that may be considered personal data. The EU has stated that personal data, once anonymised, can be published for reuse as open data since anonymised data falls outside the scope of the GDPR. Others argue, however, that, in particular, non-anonymised names in public registers are of substantial public value, including for anti-corruption purposes.

Anonymisation has proven a challenging concept for open data derived from, or relating to, personal data, particularly given it may be possible in some cases to re-identify people by linking several otherwise impersonal datasets. Current open data licensing frameworks that prohibit discrimination against any particular use of data are not equipped to restrict privacy-affecting data reuse. There have also been arguments from the open research data community to suggest that the GDPR has set the standard of consent required for data processing too high and that, in the case of specific research purposes, a lower standard of consent should be set.

The EU has stated that the PSI Directive is compliant with the GDPR and suggests that public sector bodies conduct a triple assessment to determine if personal data can be published. This includes determining if datasets contain personal data, identifying whether national access regimes restrict access to the PSI, and ensuring that PSI containing personal data can only be processed in compliance with data protection law. The GDPR also requires purpose limitation, which can cause problems with open licence terms. Purpose limitation means that personal data cannot be used for purposes other than the one for which the data was originally collected, whereas open licensing terms encourage any lawful use purpose. As of late 2018, these tensions remain unresolved with some institutions arguing for experimentation with different decision-making and data governance approaches. We anticipate that the outcome of these debates could significantly impact the discourse as to whether privacy protection outweighs the public benefits of publishing data openly in cases where the names of natural persons are relevant for particular purposes, such as fighting corruption. This discussion will be particularly interesting with regard to high-value datasets as some datasets are argued to be high-value if they are highly disaggregated and able to link monetary transactions to individuals (e.g. spending records or beneficial ownership data).
European open data infrastructure

Beyond legal considerations, the establishment of an open, usable, and interoperable data infrastructure has been key to the development of open data in Europe. A data infrastructure includes data portals, registries, and websites, data files that meet standard formats, as well as the organisational skills and work routines that surround them.

A number of governments have linked their open data programmes to the creation of unified reference databases and registers. For example, Denmark’s Danish Data Program has integrated formerly dispersed address registers, responding to the need to restructure data, as well as to share it, in order to break down departmental silos. Like PSI, the concept of National Information Infrastructures pre-dates open data, but open data agendas have helped widen engagement with the task of creating such infrastructures, as well as creating space for a wider number of stakeholders to get involved.

The most direct element of data infrastructure created by open data agendas, however, has been the data catalogue. By 2016, all EU member states had established national data catalogues, providing central access points for government data. Via web forms or application programming interfaces (APIs), government agencies can upload data to these catalogues which index their metadata, making them searchable. The European Data Portal aggregates data from many national portals to provide Europe-wide search capacity.

In order to support this kind of information exchange across governments, a shared vocabulary was necessary. The DCAT Application Profile (DCAT-AP), a Resource Description Framework (RDF) vocabulary designed to facilitate interoperability between data catalogues was developed as a common metadata standard for describing public sector datasets in Europe. It is not only used by national portals, but some city ones as well. In 2017, only Denmark, Estonia, Hungary, Lithuania, Malta, and Portugal had not yet implemented a standardised approach to collect metadata from other portals in the country and only Belgium guaranteed that 100% of its data is uploaded automatically.

For all the work that has taken place on data catalogues, the Open Data Barometer and Global Open Data Index both suggest that much of government data remains hard to find, is poorly indexed, or is of poor structural quality. Data may be hosted on ministerial websites, often hidden multiple clicks down and missing metadata. But metadata issues are also found on the European Data Portal (as indicated by the use of non-standard metadata for licences). Data standards are one way to address issues of data structure, representation, and interoperability; however, beyond the geographic datasets released through the INSPIRE directive, there is little evidence of the adoption of data standards for open datasets. Budget and finance datasets, for example, are notable for their different nomenclature and structure. Open data and catalogue standards have made it easier to discover data across borders, but, once discovered, those datasets may lack interoperability. It is notable that while the EU has substantial programmes of work on interoperability under the European Interoperability Framework, they are not yet having a major impact on the interoperability of open datasets.
The open data ecosystem

The last two sections have outlined how open data has developed at the EU and national government level. However, the open data ecosystem in Europe includes many more actors, including an active civil society, a dynamic private sector, and an engaged research community, as well as actors working on open data at the city level. This section will take stock of recent developments in this wider ecosystem.

Cities

A large number of European cities have developed local open data initiatives, and many applications of open data have been developed at the city level. Cities and their public institutions publish a variety of data relevant for urban planning, tourism, transport, mobility, green transition, utilities, and energy management. Several trailblazers should be mentioned, including European capital cities such as Helsinki, Amsterdam, Berlin, Copenhagen, London, Paris, Stockholm, Vienna, Lisbon, Dublin, and Vilnius. Other important larger and medium-sized cities which have developed open data policies as well as portals include Barcelona, Florence, Gdansk, Ghent, and Thessaloniki.

The vision and strategies of these cities may differ, but open data policies are frequently developed as part of broader digitisation and smart city strategies. As a recent report of the European Open Data Portal notes, often these strategies are developed top-down by municipal governments. Beyond merely providing open data portals, several cities have organised awareness-raising workshops, skills training, and hackathons. A number of cities organise recurring events and meetups, such as Data Dive Ghent and Helsinki Loves Developers, or collaborate with local schools to include digital skills training in curricula. These events are used as opportunities to discuss the usefulness of data or to consult with citizens about their data needs. Independent civil society initiatives, such as Open Data Manchester, city-level chapters of the Code for Europe network, and chapters and affiliates of the Open Knowledge network have also played a role in organising meetups and hackathons, as well as building local networks around open data.

The European Commission plays a role in supporting the use of open data in cities. As our analysis of Horizon 2020-funded projects (see above) shows, the European Commission has given support to several city-related initiatives, including activities to improve mobility, as well as smart city initiatives like CITYKEYS and CitySDK, which have explored the possibility of creating interoperable data infrastructures at the city level. A number of European initiatives are evident in Chapter 16: Urban development.

Despite advances in European cities, there are still several open questions and challenges, not least of which concerns who gets to participate in the opportunities created by open data. The European Data Portal reports that a “lack of awareness plays a part on both the side of the data publisher and the data user, as data publishers are not always aware of the relevance and potential of Open Data, and users are not always aware of the Open Data available”. Although Europe
ranks high in terms of education levels and technical skills, a look at open data use for citizen participation in Milton Keynes (UK) found that the majority of citizens lack key data literacies, and that building these literacies is a long-term challenge. Further research is needed to more fully investigate how well current engagement, awareness raising, and capacity-building activities are designed. Beyond technical skills to use data, new literacy skills should incorporate the politics of data and the institutional embeddedness of data, foster public pathways for engaging with how open data is produced and published, and be complemented with tangential skills from the political and social sciences. Crucially, future research must also ask who benefits? For example, are women engaged by hackathon approaches?; are skills training opportunities addressing the needs of marginalised communities?; or are programmes being designed to address specific problems in communities?

**Enterprise and private sector**

Several studies have been carried out to understand the economic benefits of open data. These studies look into the revenue generated from open data, the types of data used by enterprise, and gains from using open data. A 2015 study by the ODI identified and analysed 270 UK-based companies using open data and reported on their annual turnover and workforce size. Another Europe-wide study estimated the effects of open data on job creation, market size, cost savings, and efficiency gains, suggesting that open data will generate a total turnover of €59.7 billion. The study noted differences in market maturity with the UK, France, Germany, and Spain currently leading.

There is, however, a recognition that developing a market around open data may need some government support. From 2014 to 2017, the EU funded the Open Data Incubation Network (ODINE), supporting 57 companies in 18 countries. ODINE operated as a six-month incubation programme, funding companies with up to €100,000 and providing mentoring, business, and data training, as well as visibility at international events and introductions to investors. Another initiative with co-funding from the EU was FINODEX, a virtual accelerator supporting SMEs and web entrepreneurs to develop products and services that build upon Future Internet Fiware technologies and the reuse of open data. The UK-based ODI also runs its own startup incubation programme. It should be noted, however, that many of the potential economic benefits of open data come not from startups but from improvements to existing business practices.

While Europe does have reasonably good private sector engagement with open data, market barriers and market competition across companies and grant-funded networks may limit the full development of open data businesses. For example, a commonly noted problem is that public procurement is not conducive to supporting lean startups that do not meet revenue requirements or other procurement rules. In addition, organisations using open data might struggle to raise capital, depending instead on grants, sponsorships, membership contributions, volunteerism, and individual resources. Future research could go deeper in investigating how government procurement patterns affect open data startups and in identifying funding needs for open data businesses, including how existing funding initiatives support social enterprise, startups, and SMEs.
Non-profits and civil society

Europe’s open data ecosystem has benefited from well-developed civil society and non-profit networks working to bring together national or regional interest groups for open data. At the local and national level, groups like Open Data Manchester, Open Knowledge Belgium, and Open Knowledge France are examples of the kind of grassroots communities that came together around open data a decade ago and have developed a range of programmes and organisational capacities to advocate for, support, and work with open data. Groups like these are brought together through networks, including the Open Knowledge network, the ODI nodes, and Creative Commons. Other Europe-wide networks, such as Code for Europe, also convene events relating to open data and connect groups across countries. In addition, policy networks, such as the European Thematic Network on the Digital Public Domain (Communia), have been active on open data issues. The European Commission helps to spark cross-country collaboration by funding consortia through its Horizon 2020 programme.

Interviews with members of the Open Knowledge network suggest that civil society organisations (CSOs) have often played a crucial role in establishing a culture of openness within national governments. CSOs have helped in the exchange of best practices between public bodies, business, and individuals. For example, local Creative Commons chapters helped to promote open licensing and to influence open data policies in a number of countries. Organisations like Communia continue to engage with copyright and licensing issues. Over the last ten years, several of these networks have seen a rise and fall. OKI, in particular, experienced a period of significant growth and activity between 2010 and 2014, organising a number of key regional summits and thematic working groups, but many have since become inactive. The Open Education Working Group remains as one of the active working groups in Europe (see Chapter 6: Education).

One key challenge for civil society has been accessing sustainable funding to work on open data as opposed to short-term project funding. Membership contributions and individual donations are used by some networks, such as the Swiss Open Knowledge Chapter, and individual organisations, such as Civio in Spain, to fund activities, although, in general, project grants remain the major source of funding for open data-related CSOs.

Research and development

The EU produces a substantial share of all open data-related research. According to a systematic review of open data research produced in 2016, universities like Delft University in the Netherlands, the University of the Aegean in Greece, the University of Southampton in England, or the Copenhagen Business School in Denmark, produce a large portion of all open data articles reviewed. Most studies in this review also chose activities in the EU region as their object of study.

The review suggests that European research covers a fairly wide range of topics, including open data policy design in Europe, the impact of European open data policies, the use of open
data for policy-making in Europe, analyses of public data infrastructures, business models for PSI reuse, or technical descriptions of European data portals and other elements of data infrastructure. Open data research has contributed theoretical, as well as empirical, case studies of European open data initiatives, drawing on several disciplines and intellectual traditions, such as information management, critical data studies, science and technology studies, new media studies, and others.

Institutionally, several universities are dedicating research programmes to the study of open data, such as the University of Southampton, Delft University, or Ghent University. Research institutions are also involved in a number of collaborations with both non-profit and private sector organisations, such as the participation of Fraunhofer Fokus and University of Southampton as part of the consortium that has developed the European Data Portal.

**A thriving ecosystem?**

It is challenging to assess whether Europe’s open data ecosystem is thriving, stalling, or receding. The history of the open data ecosystem shows volatility. A notable boom for civil society and open data communities took place between 2010 and 2013. However, since then, some communities have continued to thrive, while other national and local communities have become less active or have disappeared completely. This may not necessarily indicate a weaker ecosystem as some anecdotal evidence suggests there may have been a spillover effect of expertise transitioning from volunteer-based organisations to government and business. Along these lines, weakening civil society networks might indicate a loss of engagement with open data or a shift of the priorities of network members, but could also indicate a transition from volunteerism to professional roles with the open data community. In a similar vein, it is currently unclear to what extent different models of investment in open data, such as business incubation and hackathons, contribute to the sustainable growth of open data-related organisations.

Our conversations with the community indicate that volunteerism is a major challenge in open data and may not be a sustainable organising principle for the open data sector for much longer. There are no figures available on how much labour is provided unpaid, but interviews with several organisations suggest that continued reliance on volunteers makes it challenging for organisations to sustain themselves or to scale activities. This is especially true when volunteers rarely play a leadership role, face conflicting commitments, and have sometimes been active in their roles for more than a decade, pointing to the fact that the renewal of open data communities in Europe may be overdue.

The real measure of the effectiveness of the open data ecosystem in Europe must be its impact. While there is evidence of impact and return on investment, it is uneven across political, economic, and social domains. As the EU Open Data Maturity Report states, governments still struggle to evaluate the social impact of their open data initiatives. This highlights the need for future work to improve impact measurement with a particular view to influencing strategic government investment in open data initiatives.
Conclusion

Europe played a key role in shaping the open data space a decade ago. This was demonstrated by the early adoption of rules for PSI release complementing the rise of open data, by the development of key infrastructures used for open data around the world (e.g. CKAN), by Europe’s global influence on open licensing, and by the promotion of international open data principles through the G8 Open Data Charter. Significant progress has been made in the development of open data policies in almost all countries, as well as in the development of catalogues at the EU and national levels with metadata standards for the exchange of information across these catalogues.

However, the coordination of policies across different administrative levels remains a political and administrative challenge, and many jurisdictions develop their own policies with different requirements regarding access conditions, data management, licensing, and the use of data standards. The INSPIRE Directive was a key achievement for a geospatial data standard, but for other areas, such as fiscal data, harmonisation efforts are only just starting. Standard open licences are widely used, but governments continue to apply bespoke licences, possibly creating data use silos that act against the creation of a DSM. Data markets also still interpret access to data differently, and controversy persists around whether the release of some data as open will distort competition.

The data ecosystem seems to develop at different paces in some countries with France and the UK being more advanced than others. Open data networks in civil society are volatile and may thrive moving forward or they may disintegrate. Interpreting this development, or how much positive spillover is taking place as individuals move from voluntary to professional engagement with open data, requires further study. Additionally, research into the sustainability of open data enterprises needs to pay more attention to market barriers and competition among open data organisations.

A decade on, a lack of awareness and skills to use data is still seen by governments as a major impediment to impact, illustrating the need for governments to continue to invest in educational programmes and training. Finally, it remains unclear to what extent the open data ecosystem includes marginalised groups or develops solutions with and for them, and research to date indicates that the impact of open data across the region is unevenly distributed toward organisational efficiency gains and less toward social goals. This should give us pause and inspire a new set of objectives for future open data activities that place a greater focus on inclusion and social impact.
Further reading


About the authors

Rufus Pollock is an entrepreneur, researcher, and technologist working to create an open information age. He is founder and President of Open Knowledge International (OKI), a leading international non-profit that empowers people and organisations with access to information and the tools and skills to make sense of it. Rufus is a recognised global expert on open data and open knowledge. He has worked with multiple governments, CSOs, the World Bank, the United Nations, and the private sector. You can follow Rufus at http://rufuspollock.com/ and learn more about OKI at https://www.offn.org.

Danny Lämmerhirt is Research Coordinator at OKI. His work focuses on the politics of data, metrics, policy, data ethnography, organisational theory of open systems, and the data commons. Danny leads work on the Global Open Data Index and co-chairs the Open Data Charter’s Measurement and Accountability Working Group. You can follow Danny at https://www.twitter.com/danlammerhirt.
How to cite this chapter


This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) licence. It was carried out with the aid of a grant from the International Development Research Centre, Ottawa, Canada.

Endnotes


10 Ibid.


15 https://www.economie.gouv.fr/republique-numerique-ouverture-donnees-d-interet-general
17 https://inspire.ec.europa.eu/index.cfm/pageid/2/list/7
18 http://inspire-geoportal.ec.europa.eu/
24 Ibid.
26 http://data.gov.be/fr/conditions-dutilisation
36 See http://opendefinition.org/licenses/odc-odbl/ and http://opendefinition.org/licenses/odc-by/
37 See also https://creativecommons.org/share-your-work/licensing-considerations/version4/
39 https://www.europeandataportal.eu/en/content/show-license

EUGDPR.org. (2018). The EU General Data Protection Regulation (GDPR) is the most important change in data privacy regulation in 20 years. https://eugdpr.org/


https://www.europeandataportal.eu


Ibid.


Beyond focusing on technical skills to enhance data literacy, we suggest that a more profound engagement with open data requires what could be called “data infrastructure literacy”. This form of literacy inquires how data is being made, the institutional embeddedness and politics of data, and how data is used institutionally for a variety of purposes, including public policy. See Gray, J., Gerlitz, C., & Bounegru, L. (2018). Data infrastructure literacy. *Big Data & Society, 5*(2). https://journals.sagepub.com/doi/10.1177/2053951718786316


Ibid.

University of Southampton. (2015). Leading the open data revolution. https://www.ecs.soton.ac.uk/research/leading_the_open_data_revolution

