EUROPEAN COMMISSION

Brussels, XXX
[…](2011) XXX draft

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

Open data
An engine for innovation, growth and transparent governance

COMMUNICATION FROM THE COMMISSION TO THE EUROPEAN PARLIAMENT, THE COUNCIL, THE EUROPEAN ECONOMIC AND SOCIAL COMMITTEE AND THE COMMITTEE OF THE REGIONS

Open data
An engine for innovation, growth and transparent governance

1. INTRODUCTION

The central aim of the EU 2020 strategy is to put Europe's economies onto a high and sustainable growth path. To this end, Europe will have to strengthen its innovative potential and use its resources in the best possible way.

One of these resources is public data[1] — all the information that public bodies in the European Union produce, collect or pay for. Examples are geographical information, statistics, weather data, data from publicly funded research projects, and digitised books from libraries. This information has a significant — currently untapped — potential for re-use in new products and services and for efficiency gains in administrations. Overall economic gains from opening up this resource could amount to €40 billion a year in the EU. Opening up public data will also foster the participation of citizens in political and social life and contribute to policy areas such as the environment.

This Communication presents a package of measures to overcome existing barriers and fragmentation across the EU, as part of the Digital Agenda for Europe. It consists of three strands that reinforce each other:

- Adapting the legal framework for data re-use. A proposal for a revised Directive on the re-use of public sector information and a revised Commission Decision on the re-use of its own information are adopted together with this Communication,

- Mobilising financing instruments in support of open data, and deployment actions such as the creation of European data-portals,

- Facilitating coordination and experience sharing across the Member States.

The actions focus on areas where the functioning of the internal market is at stake and where common standards and approaches will lead to new and better services and information products for the European consumer. They build on and do not affect the national regimes for access to information.

[1] The concepts of 'data' and 'information' as used in this Communication are interchangeable and refer to any content, whatever its medium.

2. OPEN DATA, WHY DOES IT MATTER FOR EUROPE?

2.1. Untapped business and economic opportunities

The 2009 Digital Britain Report described data as '*an innovation currency'* and *'the lifeblood of the knowledge economy*'[2]. A large part of this innovation currency is produced, collected or paid for by governments across the EU. It is an essential raw material for a wide range of new information products and services that build on new possibilities to analyse and visualise data from different sources.

> *These products range from decision support systems for businesses, location-based services and car navigation systems to weather forecasts and other 'apps' for our smartphones.*

The market size and growth of the geographic information sector shows the potential of public data as an engine for job creation. The German market for geo-information in 2007 was estimated at € 1.4 billion, a 50 % increase since 2000[3]. In the Netherlands, the geo-sector accounted for 15 000 full time employees in 2008. Other areas such as meteorological data, legal information and business information also form the basis of steadily growing markets.

A recent study estimates the total market for public sector information in 2008 at € 28 billion across the EU[4]. The same study indicates that the overall economic gains from further opening up public sector information by allowing easy access are around € 40 billion a year for the EU27. The total direct and indirect economic gains from PSI applications and use across the whole EU27 economy would be in the order of € 140 billion annually.

2.2. Addressing societal challenges

Intelligent processing of data is essential for addressing societal challenges. Data can for example be used to enhance the sustainability of national health care systems. The 2011 McKinsey report estimated that effective use of data could generate $ 300 billion in value per year[5].

> *Progress in genomics, drug discoveries and the diagnosis and treatment of serious diseases such as cancer or heart failure increasingly depend on sophisticated data capturing and analysis techniques.*

Data management is also essential for tackling *environmental* challenges. Examples are the processing of energy consumption patterns to improve energy efficiency or of pollution data in traffic management. Informed policy decisions in the areas of transport, land use and climate change depend increasingly on analysis of the available data.

---

2 http://www.official-documents.gov.uk/document/cm76/7650/7650.pdf.

3 Assessment of the Re-use of Public Sector Information in the Geographical Information, Meteorological Information and Legal Information sectors, MICUS, December 2008.

4 Review of recent studies on PSI re-use and related market developments, G. Vickery, August 2011.

5 Mc Kinsey report on 'Big Data, The next frontier for innovation, competition and productivity', May 2011.

2.3. Accelerating scientific progress

Scientific activities are increasingly undertaken through global collaboration on the internet, using very large data collections, huge computing resources and high-performance visualisation. **e-science** (research enabled by e-infrastructure/ICT) is essential for meeting the challenges of the 21st century in scientific discovery and learning. The data used come from simulations, digital instruments, sensor nets, and observatories.

The 2010 report 'Riding the Wave'[6] underlined the crucial role of data for science and its potential to change the very nature of scientific process.

> *Wide access to scientific data will for example help researchers in different domains to collaborate on the same data set, to engage in entirely new forms of scientific research and to explore correlations between research results.*

The shift in the scientific process brought about by e-science will increase research productivity and prompt new and unexpected solutions to societal challenges. Furthermore, the cross-fertilisation between publicly funded research and the commercial sector in the 'online European Research Area' will increase the pace and impact of innovation.

2.4. Need to act at all levels: local, regional national and EU level

Public data are produced at all levels of government. The conditions under which the data are made available for commercial or non-commercial re-use have an effect on competition and competitiveness.

Many of the products and services based on public sector information have a cross-border nature. You do not want your mobile service to stop at the border, and Europe-wide business information services with gaps for one or more countries will lose much of their interest. Moreover, in order to benefit from the size of the internal market and to help SMEs grow beyond their national borders, the deployment of services across the whole EU should be facilitated.

Therefore, the same basic conditions should apply across the European Union. This will create a level playing field and stimulate a thriving market of innovative products and services based on public data.

3. WHERE DO WE STAND AND WHAT HAS BEEN ACHIEVED SO FAR?

3.1. Where do we stand?

The thrust towards open data is gaining momentum in several Member States. They are embracing the concept for reasons of transparency, administrative efficiency and the economic potential of re-use. They are supporting open government through legislation and practical measures, such as the production of data in machine-readable formats and the creation of data portals.

> *The UK has set up the data.gov.uk portal, which brings together data from government organisations at all levels. Other Member States are creating similar*

[6] Final report of the High Level Expert Group on Scientific Data, October 2010.

portals, e.g. France with ETALAB. Data portals also exist at regional level, such as dadesobertes.gencat.cat in Catalunya and dati.piemonte.it in Piemonte, Italy.

Nevertheless, the degree of initiative and the awareness of open data issues are uneven among the Member States. There is a risk that Europe will miss out on the opportunities offered by open data, and will lag behind other regions where open data policies are well established.

3.2. What has been achieved at EU level?

3.2.1. Legislative framework and policy initiatives

(1) Directive on the re-use of public sector information

The 2003 Directive on the re-use of public sector information[7] set out the general legislative framework at European level. The Directive provides for a minimum degree of harmonisation. The 2009 review of the Directive indicated that, in spite of progress since 2003, barriers to the cross-border use of public sector information still existed. Some of these barriers can be tackled within the existing legislation, others cannot.

(2) Policy initiatives

The general re-use policy is complemented by legislative or policy initiatives in specific sectors. Examples are:

- the Access to Environmental Information and INSPIRE Directives[8], aimed at the widest possible dissemination of **environmental information** and the harmonisation of key datasets;
- the Commission Communication on Marine Knowledge 2020[9], aiming amongst other things to make **marine data** easier and less costly to use;
- the initiatives within the 2008 **Action Plan**[10] **for the Deployment of Intelligent Transport Systems (ITS)**, looking at, amongst other things, access for private service providers to travel and real-time traffic information;
- the Commission's policy on open access to **scientific information**[11], which includes a pilot for open access to publications resulting from projects funded by the European Union and a pan-European, participatory **e-Infrastructure of Open Access Repositories**; The JRC publications repository is also relevant in this context.
- the policies for the digitisation of **cultural heritage** and the development of Europeana, Europe's digital library, archive and museum, aiming to ensure the widest possible use of digitised cultural material and the related metadata.

---

7 OJ L 345/90, 31.12.2003.

8 Directives 2003/4/EC, OJ L 41/26, 14.2.03 and 2007/2/EC, OJ L108/1, 14.3.2007.

9 COM(2010) 461 final.

10 COM(2008) 886 final/2.

11 Communication on scientific information in the digital age, COM(2007) 56, Communication on ICT infrastructures for e-Science COM(2009) 108 and related Council Conclusions.

The re-use of public sector data, whether for commercial or non-commercial purposes, should fully respect European and national **privacy legislation**. The objectives of fostering open government data and protecting personal data can reinforce each other if based on pro-active and conscious information management by the public sector.

Moreover, the implementation of open data policies should fully respect the **intellectual property rights** of third parties and the European Union's obligations under international treaties on intellectual property rights.

3.2.2. Co-funding of R&D&I

The Commission has supported open data through its funding programmes, in particular the Framework Programmes for Research and Development, the Competitiveness and Innovation Programme, and the ISA programme. The projects cover a range of research and application areas and types of organisations.

> *The **Linked Open Data (LOD2)** project[12] started in September 2010 and will run for 4 years. It addresses exploitation of the web as a platform for data and information integration, and the use of semantic technologies to make government data more useable.*
>
> *The **OpenAIRE** project[13], which started in December 2009 with partners from 25 EU countries and several associated countries, aims to build a participatory infrastructure for the EC Pilot for Open Access to Research Information.*
>
> *The ISA Action on semantic interoperability (SEMIC.EU)[14] promotes the idea of Open Government Metadata as a first step towards metadata alignment at both national and European level.*

3.2.3. Practising what is preached

In 2006, the Commission put in place an open re-use policy for its own information resources through the Decision on the Re-use of Commission Information. It makes all generally accessible Commission information available for commercial and non-commercial purposes. Examples are Eurostat data and translation memories of EU institutions. The information is normally available for free, or in exceptional cases in return for the marginal costs of dissemination.

4. CHALLENGES AND OPPORTUNITIES

4.1. New opportunities through technological progress

Several developments spur new opportunities for the re-use of data — including public sector data — in new information products and services.

---

12 http://lod2.eu/.

13 http://www.openaire.eu/.

14 http://www.semic.eu/

First of all, the *amount of data* in the world is exploding. The US Library of Congress collected 235 terabytes of data in April 2011 alone[15]. Growth in global data generated per year is projected at 40%.

Part of this growth comes from *new types of data*. Already, more than 30 million networked sensor nodes can be found in the transport, automotive, utility and other sectors[16].

In parallel, we are witnessing a revolution in the *technologies* for analysing, exploiting and processing data, for example on multilingual search and the automatic extraction of meaning from networks of sensors.

4.2. Challenges and remaining barriers

In recent years some progress has been made to open up public data, but different barriers persist.

The legal framework

Despite the minimum harmonisation in 2003 through the Directive on the re-use of public-sector information, significant **differences in national rules and practices** persist. This leads to fragmentation of the internal information market and hinders the creation of cross-border information services.

Differences are the clearest with respect to charging, with cost recovery practised in some cases and re-use free or practically free in others. An analysis of recent studies indicates that from a macro-economic point of view the open data model leads to better overall results[17]. A series of case studies on public sector bodies that moved from full cost recovery to a marginal costs system show that the move not only increased re-use, but also benefited the public sector bodies concerned[18]. Moreover the open data approach eliminates possible monopolistic tendencies based on single-source data.

Awareness of public organisations and businesses

Another key factor is the **lack of awareness among public organisations** of the potential of open data. There is still a widespread fear of losing control. Some of the concerns are legitimate, such as privacy protection, national security and the need to protect the intellectual property rights of third parties. Other arguments seem to be rather excuses for inaction.[19]

Changing the mindset in administrations requires strong political commitment at the highest level and a dynamic dialogue between stakeholders, including administrations and public data holders, businesses and the academic community. Pilot and test cases, the sharing of good practices, and mobilisation campaigns (using for example open data competitions) can help the public sector in adopting a culture of open data. This will also increase business awareness of data availability and the opportunities it offers.

---

15 McKinsey, op.cit.
16 idem.
17 G. Vickery, op.cit.
18 Study on 'Pricing of Public Sector Information', Deloitte consulting and others, June 2011.
19 Some examples at http://www.dr0i.de/lib/2011/07/04/a_sample_of_data_hugging_excuses.html

Practical and technical hurdles

Finally, there are **practical and technical** issues to be addressed. Language barriers and interoperability aspects need to be tackled so that information resources from different organisations and countries can be combined. The availability of the information in a machine-readable format as well as a thin layer of commonly agreed metadata could facilitate data cross-reference and interoperability and therefore considerably enhance its value for re-use. And the technical infrastructure needs to be in place to ensure the availability of information in the long term. In addition, more support is needed for R&D and innovation in data analysis and visualisation tools.

4.3. Conclusion

Public data exploitation holds enormous potential for the EU economy and consumer welfare. However, the existing regulatory tools and their implementation, the lack of awareness of administrations and businesses and the slow uptake of innovative technologies are holding back the development of a true market for the re-use of public data and do not allow the maximum benefits to be reaped from the new opportunities that data and evolving technologies offer.

These considerations have led the Commission to revise and strengthen its public data strategy by targeting the legal framework for re-use and available support tools.

5. ACTIONS AT EUROPEAN LEVEL: REVISED STRATEGY FOR PUBLIC OPEN DATA

The Commission's revised strategy is built on three complementary strands:

- Adapting the framework for data re-use, including legal, soft law and policy measures,
- Mobilising financing instruments by prioritising open data in R&D&I and infrastructure programmes,
- Facilitating coordination and experience sharing across the Member States.

5.1. A revised legislative framework conducive to openness and re-use

This Communication is accompanied by a proposal for revising the Directive on the re-use of public sector information, by:

– Introducing the principle that all public information that is not explicitly covered by one of the exceptions is re-usable for commercial and non-commercial purposes;

– Setting the amount that can be charged for public sector information at not more than the marginal costs of dissemination; in exceptional cases, charging the full costs for producing and disseminating the information remains possible.

– Expanding the scope of the Directive to include libraries, archives, museums and university libraries, though in a way that limits the possible financial effects and does not impose a major administrative burden on these institutions.

Also for other public sector bodies the modifications in the Directive will entail at most a limited financial and administrative burden that is by far outweighed by the expected benefits for society as a whole.

At the same time, the Commission is updating the rules on the re-use of its own information resources. The main changes concern the inclusion of the research information produced by the Joint Research Centre, measures to improve the implementation of the Decision, and a provision on machine-readable formats. In 2012 the Commission will explore whether and how similar rules could be taken up by the other EU institutions and key European agencies.

In its forthcoming legislative and policy initiatives in areas such as transport, environment, maritime policy and space policy, the Commission will take full account of the opportunities offered by open data. Where relevant, it will go beyond the minimum conditions set by the Directive on the re-use of public sector information, to ensure the widest possible use of data resources.

5.2. Soft law for open research information

Because of the specificities of research data, the Commission will set out in detail and in separate documents its strategy for *scientific and research data* and associated infrastructures. It intends to adopt in 2012 a Communication and Recommendation on the accessibility and preservation of scientific information. It will work with Member States to step up their activities to provide open access to scientific information on the basis of a concrete set of measures. In parallel, the Commission will detail how it will deal with the results of research funded by the European Union..

5.3. Financing and support measures: Stimulating R&D&I in open data

The Commission will continue stimulating activities to open up government data through its funding programmes. It will use different instruments to stimulate the market, to test and promote the development of innovative solutions, and to ensure the widest possible uptake of open data.

(1) Support for R&D&I

The Commission will continue to support *R&D* in data-handling technologies, e.g. data mining, analytics or visualisation. In the period 2011-2013 the Commission will spend around € 100 million on R&D in these fields. Information management is also one of the priority areas envisaged for ICT in Horizon 2020, which will cover EU support for research and innovation in the period 2014-2020.

The Commission will support *technology innovation and uptake* through pilot actions, testing and showcasing innovative applications such as geographical information systems and location-based services (GIS) and creative content applications in education, culture or fashion. These will be supported under the CIP-ICT PSP in 2012-13 and further support is envisaged under Horizon 2020.

In addition, the Commission will help organise *open data competitions* to foster the development of new information services,, and take initiatives to improve *access to capital* for entrepreneurs who build new information services based on public sector data.

(2) Support for data infrastructures — data portals for Europe

In order to facilitate the development of information products and services combining data from across the European Union, the Commission will work towards the creation of two interlinked European data portals.

In 2012 it will start operating a portal that will make the Commission's own data resources and those of other European institutions and agencies easily accessible and usable. In parallel, metadata definitions with high re-use potential will be identified and promoted. Standard conditions for use will aim to ensure the widest possible uptake of European data in innovative information services, in line with the Commission Decision on the re-use of Commission information.

The Commission will work together with Member States, public sector bodies and regional aggregators to establish a pan-European data portal that will start operating in 2013, and that will give direct access to a range of datasets from across the EU (including the data available through the Commission portal). Where possible, the work will build on existing structures and developments, and the portal will gradually expand its content to cover key datasets from all the Member States.

Support will be provided from the CIP programme in the inception phase (2011-2013). In the period 2014-2020, funding for the European e-service infrastructure for public data will come from the Connecting Europe Facility[20].

The Commission will also continue to support digitisation activities and the development of the Europeana platform, which addresses an important aspect of data re-use.

(3) Support for research data infrastructures

The Commission supports in FP7, and envisages continued support in Horizon 2020 for the development of a persistent and robust service infrastructure for scientific data in Europe that responds to the needs of the data-intensive science and research of 2020, guided by the recommendations of the 'Riding the Wave' report. It will allow access to and interaction with a continuum of information, from raw observational and experimental data to publications in all areas of science.

This infrastructure encompasses technical, organisational and regulatory aspects that require extensive coordination with Member States and also with third countries and international organisations to ensure global interoperability and reciprocal access. The Commission will work together with our international partners to develop standards for global data access and interoperability.

5.4. Coordinating measures at Member State and EU level

The Commission will continue facilitating coordination and experience sharing across the Member States, in particular through:

- The PSI group, a Member States' expert group for the exchange of good practices and initiatives supporting public-sector information re-use,

[20] See the Commission proposals for the CEF, COM(2011)665, COM(2011)657/3.

- The Public Sector Information platform. This web portal provides news on European developments, good practices, examples of new products and services, and legal cases concerning PSI re-use,

- The LAPSI network, which analyses legal issues related to public sector information and fosters debate among researchers and stakeholders. It will produce a set of guidelines for access and re-use policies and practices,

- The ISA action on semantic interoperability.

The Commission will continue to support and participate in policy advisory groups such as the e-Infrastructures Policy Forum and the e-Infrastructures Reflection Group, important for coordination between Member States on scientific data infrastructures.

6. CONCLUSION

Information produced, collected or paid for by public organisations across the European Union is a key resource in the information economy. At the moment, its full potential is far from being realised. In this Communication, the Commission proposes concrete steps to unlock the potential of Europe's public sector resources, ranging from a review of the Directive on the re-use of public sector information to the creation of a pan-European portal.

Member States can contribute to making open data a reality through the rapid adoption, transposition and implementation of the revised Directive on the re-use of public sector information. This will create the conditions for economic activity based on open data, and will stimulate cross-border applications.

In addition, Member States should formulate and implement open data policies, taking up good-practice examples from across the EU. Support should for example be given to open data pilots and open data competitions, in particular those targeting the development of cross-border products and services.

Finally, the Commission calls on the Member States to contribute to the development of the pan-European data portal. The Commission will engage in discussions with experts from the Member States in order to ensure that the portal has a solid foundation. During the deployment phase, Member States will have to make a wide range of datasets available through the portal to turn it into a successful service and a basis for economic activity.

The Commission invites the European Parliament and the Council, within their respective responsibilities, to create the right framework conditions for the re-use of public sector information across the European Union, and to support the projects and infrastructures that can turn Europe’s public data into a motor for innovation, growth and transparency.

Open data strategy, key measures

The legislative framework

- Proposal for a revised Directive on the re-use of public sector information, December 2011;
- Revised Commission Decision on the re-use of Commission information, December 2011; Work to expand the regime to other European Institutions and Agencies, 2012;
- Open data to be taken up in sector-based legislative and policy initiatives.

European open data portals and platforms

- Portal giving access to Commission data and data from other EU institutions and agencies, spring 2012;
- Launch of a pan-European data portal, giving access to datasets from across the EU, spring 2013, following preparatory work with Member States from 2011;
- Co-funding of the European e-service infrastructure for open data through the Connecting Europe Facility 2014-2020.

Open data for science

- Communication and Recommendation to the Member States on scientific information, early 2012;
- Expansion of the open access pilot for scientific publications to the whole of Horizon 2020 + pilot with open access to research data.

Research and innovation

- Research and innovation projects relevant for open data, in particular through FP7, CIP and Horizon 2020, with funding for research infrastructures supporting open access to research articles and data;
- Open data competitions (2012-2013) + improving access to capital for entrepreneurs in this area.

Targets

The Commission is committed to:

- the launch of a Commission open data portal in 2012.
- the launch in 2013 of a pilot-portal with a multilingual interface and search facilities, with datasets from across the EU.

The work with the Member States should lead to:

- the formulation and implementation of open public data policies in all Member States by early 2013.
- 1/3 of all available structured government data in the Member States searchable through the pan-European data-portal by 2015.

Overall envisaged impact:

- by 2017 (three years after the expected transposition date of the Directive on the re-use of public sector information), the overall gains of PSI re-use to reach € 100 bn per year in the EU, including new business development and efficiency gains in public sector services.