



Brussels, 2.7.2014
COM(2014) 442 final

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

Towards a thriving data-driven economy

{SWD(2014) 214 final}

1. Introduction

The European Council's conclusions of October 2013 focused on the digital economy, innovation and services as drivers for growth and jobs. They called for EU action to provide the right framework conditions for a single market for big data and cloud computing.

This Communication responds by sketching the features of the data-driven economy of the future and setting out some operational conclusions to support and accelerate the transition towards it. It also sets out current and future activities in the field of cloud computing.¹

This Communication builds on the results of various consultations² and on relevant legislative proposals already tabled, such as on reform of the EU rules on the protection of personal data and on network and information security.³

Global context and call to action

We witness a new industrial revolution driven by digital data, computation and automation. Human activities, industrial processes and research all lead to data collection and processing on an unprecedented scale, spurring new products and services as well as new business processes and scientific methodologies.

The resulting datasets are so large and complex that it becomes difficult to process such "big data" with the current data management tools and methods. At the same time, technological advances allow for new ways to cope with these challenges. For example, cloud computing provides large scale computing as a service to the data economy in the same way as power plants supply the manufacturing industry.

Big data technology and services are expected to grow worldwide to USD 16.9 billion in 2015 at a compound annual growth rate of 40% – about seven times that of the information and communications technology (ICT) market overall. A recent study predicts that in the UK alone, the number of specialist big data staff working in larger firms will increase by more than 240% over the next five years.⁴

This global trend holds enormous potential in various fields, ranging from health, food security, climate and resource efficiency to energy, intelligent transport systems and smart cities, which Europe cannot afford to miss.

Yet the European digital economy has been slow in embracing the data revolution compared to the USA and also lacks comparable industrial capability. Research and innovation (R&I) funding on data in the EU is sub-critical and the corresponding activities are largely uncoordinated. There is a shortage of data experts able to translate technology advances into

¹ The Staff Working Document accompanying this Communication reports on the implementation of the European Cloud Computing Strategy, COM(2012) 529

² E.g. <http://ec.europa.eu/digital-agenda/en/news/trusted-cloud-europe-survey>; <https://ec.europa.eu/digital-agenda/en/content/consultation-research-data-infrastructures-framework-action>

³ COM(2012) 10 final, COM(2012) 11 final and COM(2013) 48 final

⁴ Big Data Analytics – An assessment of demand for labour and skills, 2012-2017. *e-skills UK report on behalf of SAS UK*.

concrete business opportunities. The complexity of the current legal environment together with the insufficient access to large datasets and enabling infrastructure create entry barriers to SMEs and stifle innovation.

As a result, there are fewer successful data companies in Europe than in the USA where large players have recognised the need to invest in tools, systems and new data-driven processes. However, significant new opportunities exist in a number of sectors (from health and smart factories to agriculture) where the application of these methods is still in its infancy and global dominant players have not yet emerged.

The accelerating digitisation of public services, driven by the need to modernise, cut costs and provide innovative services, opens up further opportunities to optimise data storage, transfer, processing and analysis.

At the same time, the reported use of similar technologies for surveillance purposes, by public or private actors, is liable to feed concern and reduce trust in the digital economy among individuals and organisations. The Commission has always taken such concerns very seriously. It will continue to address them by enacting effective data protection and network and information security rules, supporting secure technologies and informing the public about ways to reduce privacy and security risks. A high level of trust is essential for the data-driven economy.⁵

To be able to seize these opportunities and compete globally in the data economy, the EU must:

- support "lighthouse" data initiatives capable of improving competitiveness, quality of public services and citizen's life. "Lighthouse" initiatives maximise impact of EU-funding within strategically important economic sectors. Possible areas include the health sector (personalised medicine), integrated management of transportation and logistics for entire regions, the management of food chains by tracking food items from farm to fork, etc.;
- develop its enabling technologies, underlying infrastructures and skills, particularly to the benefit of SMEs;
- extensively share, use and develop its public data resources and research data infrastructures;
- focus public R&I on technological, legal and other bottlenecks;
- make sure that the relevant legal framework and the policies, such as on interoperability, data protection, security and IPR are data-friendly, leading to more regulatory certainty for business and creating consumer trust in data technologies;
- rapidly conclude the legislative processes on the reform of the EU data protection framework, network and information security, and support exchange and cooperation between the relevant enforcement authorities (e.g. for data protection, consumer protection and network security);

⁵ See also JOIN(2013) 1, Cybersecurity Strategy of the European Union: An Open, Safe and Secure Cyberspace, of 07/02/2013 which set "out the actions required [...] to make the EU's online environment the safest in the world", p. 3.

- accelerate the digitisation of public administration and services to increase their efficiency; and
- use public procurement to bring the results of data technologies to the market.

A coordinated action plan involving Member States and the EU can guarantee the necessary scope and scale of the required activities, such as the building of world-class connectivity, storage and supercomputing capacities for data or the identification of areas of strategic importance for the Union where breakthroughs can be made.

By building upon ongoing sectoral activities already contributing to a data-driven economy, for example in the field of multimodal travel, this Communication seeks to initiate a debate with the Parliament, Council and other stakeholders, including the network of national digital coordinators⁶ on developing such an action plan. To steer this debate, this Communication describes the characteristics of a data-driven economy and outlines a set of initial actions to help bring it about in Europe.

2. Data is at the centre of the future knowledge economy and society

The number of ways in which digital data is generated, collected, processed and used is increasing quickly. For example, manufacturers collect and process data to optimise the flow of materials and goods while new goods and services increasingly rely on embedded data analytics (e.g. collision-avoidance systems).

According to ISO/IEC 2382-1, data is "a reinterpretable representation of information in a formalized manner, suitable for communication, interpretation or processing". Data can either be created/authored by people or generated by machines/sensors, often as a "by-product". Examples: geospatial information, statistics, weather data, research data, etc.

Provided that rules on the protection of personal data, when applicable, are complied with, data, once recorded, can be re-used many times without loss of fidelity. This aggregated value generation is at the core of the data value chain concept. For example, aggregated location information of mobile phones in cars can be re-used for real-time traffic information.

The term "big data" refers to large amounts of different types of data produced with high velocity from a high number of various types of sources. Handling today's highly variable and real-time datasets requires new tools and methods, such as powerful processors, software and algorithms.⁷

In general, analysing data⁸ means better results, processes and decisions. It helps us generate new ideas or solutions or to predict future events more accurately. As technology advances, entire business sectors are being reshaped by systematically building on data analytics.⁹

⁶ To be established in accordance with the conclusions of the European Council of October 2013.

⁷ Going beyond traditional "data mining" tools designed to handle mainly low-variety, small scale and static datasets, often manually.

⁸ This can comprise both genuine "big data" and a wealth of other data sets ("small data").

⁹ Businesses that use 'data-driven decision-making' enjoy a 5-6% increase in productivity, Big Data for All: Privacy and User Control in the Age of Analytics, O. Teme/J. Polonetsky, Northwestern Journal of Technology and Intellectual Property 2012.

The term 'data-driven innovation' (DDI) refers to the capacity of businesses and public sector bodies to make use of information from improved data analytics to develop improved services and goods that facilitate everyday life of individuals and of organisations, including SMEs.¹⁰

To facilitate exploitation and reduce transaction costs, the fewer restrictions and the more harmonised the rules on data re-use, the better. Echoing the Commission's earlier open data policy¹¹, the G8's 2013 Open Data Charter incorporates the principle of 'open by default' and stresses the need to make data freely and openly re-usable both for humans and machines.

The term "open data" refers to a subset of data, namely to data made freely available for re-use to everyone for both commercial and non-commercial purposes.

The existence of datasets, be they distributed across different locations and sources, open or restricted, and possibly including personal data that needs special protection, poses new challenges for the underlying infrastructure. Data analytics requires a secure and trusted environment that enables operations across different cloud and high-performance computing (HPC)¹² infrastructures, platforms and services.

Data-driven innovation brings vast new job opportunities. However, it requires multidisciplinary teams with highly skilled specialists in data analytics, machine learning and visualisation as well as relevant legal aspects such as data ownership, licence restrictions and data protection. The training of data professionals who can perform in-depth thematic analysis, exploit machine findings, derive insight from data and use them for improved decision-making is crucial.

The EU's Horizon 2020 (H2020) and national R&I funding programmes can address relevant technical challenges: from data creation and actuation through networks, storage and communication technology to large-scale analysis, advanced software tools and cyber security. Finally, support to stimulate sector-specific entrepreneurship and innovation is important.

3. Towards a data-driven EU economy

A prominent feature of a data-driven economy will be an ecosystem of different types of players interacting in a Digital Single Market, leading to more business opportunities and an increased availability of knowledge and capital, in particular for SMEs, as well as more effectively stimulating relevant research and innovation.

A thriving data-driven economy will have the following characteristics:

3.1. Availability of good quality, reliable and interoperable datasets and enabling infrastructure

(1) *The datasets themselves*: good quality reliable and trusted data coming from large datasets, including open data (e.g. Earth observation and other geospatial data, language

¹⁰ Data-Driven Innovation – A Guide for Policymakers: Understanding and Enabling the Economic and Social Value of Data, SIIA White Paper, 2013.

¹¹ Open data. An engine for innovation, growth and transparent governance, COM(2011) 882; Directive 2013/37/EU.

¹² High-Performance Computing: Europe's place in a Global Race, COM(2012) 45.

resources, scientific data, transport data, healthcare data, financial data, digitisation of cultural assets) being widely available for new data products. No inappropriate restrictions hinder the flow of data across sectors, languages and borders in the Digital Single Market. Users have sufficient trust in the technology, the behaviour of providers and the rules governing them;

- (2) *The flexibility required to use the datasets*: standard and shared formats and protocols for gathering and processing data from different sources in a coherent and interoperable manner across sectors and vertical markets (energy, transport, environment, smart cities, retailing, security, etc.); and
- (3) *Solid infrastructures, resources and services*: open data portals and research infrastructures that support data-driven innovation, based on fast internet and the availability of large and flexible computing resources (in particular HPC, grid and cloud computing infrastructures and services, and statistical infrastructure).

3.2. *Improved framework conditions that facilitate value generation from datasets*

- (1) *An adequate skills base*: small and large companies and universities cooperate to train a sufficient number of domain experts to meet the strong demand in the labour market. This involves an effective and efficient cross-fertilisation of talent and skills bridging diverse areas; and
- (2) *Close cooperation between players*: universities/public research institutes and private partners, especially SMEs, cooperate on R&I across sectors through facilitated access to and transfer of knowledge and technology. Such public-private cooperation ensures the availability and further development of reliable and adequate algorithms, tools and methods for descriptive and predictive data analytics, data processing, simulation, visualisation, decision support and the integration of results into new products.

3.3. *A range of application areas where improved big data handling can make a difference*

- (1) *Systems*: ICT systems able to perform sensing, actuating, computing, communication embedded in physical objects, interconnected through the Internet and providing citizens and businesses with a wide range of innovative applications and services (smart connected objects); and
- (2) *Early adopters & catalysts*: public sector bodies act as 'launching customers' and intermediaries for new data services and digital goods. The public sector has a key role in the adoption of cloud computing services and other new approaches and in the creation of trust by citizens and businesses, including SMEs.

4. An action plan to bring about the data-driven economy of the future

Progress towards a thriving data-driven economy requires community building and the right framework conditions.

4.1. Community building

1. A European Public-Private Partnership on Data

In the Commission's view, strategic cooperation through a contractual Public-Private Partnership (cPPP)¹³ can play an important role in developing a data community and encouraging exchange of best practices. In line with the principles set out in H2020, the Commission considers that a sufficiently well-defined cPPP would be the most effective way to implement H2020 in this field, notably given the required scale of impact, the resources involved and the importance of a long-term commitment.

A cPPP enshrines commitments on the part of the Commission and of the industry to engage in R&I activities and constitutes a valuable discussion forum. It steers R&I activities through a Strategic Research and Innovation Agenda (SRIA), to be coordinated with Member States' agendas, focusing all relevant efforts on the most important challenges and bottlenecks, maximising efficiency and avoiding duplication.

A cPPP on data should develop incentives to share datasets between partners and mechanisms to facilitate knowledge and technology transfers. It should collaborate with academic and research institutions so that students and researchers can experiment with realistic and large datasets while promoting exchanges between data scientists, data protection and security experts.

Industry has organised itself and is preparing a proposal for such a cPPP.¹⁴ If evaluated successfully, it could be launched by the end of 2014.

2. Digital entrepreneurship and open data incubator

Recognising the high potential of digital technologies in boosting more entrepreneurial action and transforming all types of businesses in Europe, the Commission has launched a strategy to support Digital Entrepreneurship in the Union¹⁵.

In this same spirit, within the H2020 framework, an open data incubator will help SMEs set up supply chains based on data, promote open or fair access conditions to data resources, facilitate access to cloud computing, promote links to local data incubators across Europe and help SMEs obtain legal advice.

3. Developing a skills base

The Commission will design a European network of centres of competence to increase the number of skilled data professionals. This will be complemented by the recognition of new e-infrastructure professions and skills, in line with the 'Grand Coalition on Digital Skills and Jobs' initiative¹⁶.

¹³ See Article 25 of Regulation (EU) 1291/2013 establishing Horizon 2020.

¹⁴ www.bigdatavalue.eu

¹⁵ http://ec.europa.eu/enterprise/sectors/ict/digital-entrepreneurship/index_en.htm

¹⁶ <http://ec.europa.eu/digital-agenda/en/grand-coalition-digital-jobs-0>

4. Data market monitoring tool

The Commission is setting up a data market monitoring tool to measure the size and trends of the European data market. This tool will also show the relations between the different actors in the European data economy.

5. Identification of sectoral priorities for R&I

The Commission will invite stakeholders and research communities (e.g. from the health, energy, environment, social sciences and official statistics sectors) to propose "lighthouse" initiatives that may yield the greatest social and economic benefits and should attract the necessary public and private funding.

4.2. Developing framework conditions

4.2.1. Availability of data and interoperability

1. Fostering Open Data policies

To facilitate the implementation of the EU open data policy¹⁷ and legal framework¹⁸, the Commission is preparing guidelines on recommended standard licences, datasets and charging for the re-use of documents.

The Commission and other EU bodies are releasing their own documents as open data through the EU Open Data Portal. In addition, a pan-European open data digital service infrastructure under the Connecting Europe Facility programme will provide a one-stop-shop to open data across the EU.¹⁹ Measures to promote scientific discovery and collaboration across disciplinary and geographical boundaries are included in the Commission's scientific information package²⁰.

The aim of further opening up data for access and re-use is also pursued by a number of Commission initiatives covering sector-specific data (transport, environment, etc.) as well as through open access to H2020 results²¹.

2. Data handling tools and methods

In order to encourage R&I on business intelligence, decision support processes and systems supporting SMEs and web-entrepreneurs, H2020 addresses descriptive and predictive data analytics, data visualisation, artificial intelligence and decision-making software tools and algorithms.

¹⁷ COM(2011) 882 final

¹⁸ Directive 2013/37/EU

¹⁹ Regulation 2014/283/EU on guidelines for trans-European networks in the area of telecommunications infrastructure.

²⁰ Towards better access to scientific information, COM(2012) 401; Commission Recommendation on access to and preservation of scientific information, C(2012) 4890.

²¹ http://ec.europa.eu/research/participants/data/ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-pilot-guide_en.pdf

Other topics include proofs of concept and prototypes of cloud-based data infrastructure enablers (i.e. Platform as a Service and Software as a Service) for extremely large or highly heterogeneous datasets and actions to deal with large, complex and data-intensive systems and services.

Finally, H2020 will stimulate the setting-up and networking of competence centres to support SMEs in developing, accessing and taking up data technology or services in their products, business processes or other activities.

3. Supporting new open standards

Open standards and data interoperability are priorities in various Commission policies. This is reflected in ongoing initiatives to set EU-wide standards within important economic sectors, such as transport. The ISA programme²² facilitates the use of common core data standards for national administrations. To help create a climate of open data exchange, the Commission will support the mapping of existing relevant standards for a number of big data areas (e.g. smart grid, health, transport, environment, retail, manufacturing, financial services).

Future actions under H2020 will identify industrial sectors that are sufficiently homogenous in their activities to further develop relevant standards.

4.2.2. Enabling infrastructure for a data-driven economy

1. Cloud computing

The actions of the European Cloud Computing Strategy on transparency of standards, voluntary EU-wide certification, safe and fair contract terms and conditions for cloud users and the establishment of a European Cloud Partnership (ECP) facilitate a more rapid adoption of trusted cloud computing that will boost the data-driven economy.

The ECP Steering Board recently published its Trusted Cloud Europe (TCE) report²³. The results of a subsequent survey indicated consistent support for a vision of a Trusted Cloud Europe based on the rapid adoption of the general data protection regulation in Europe and effective mechanisms to ensure European "data sovereignty" in the cloud. As follow-up, the Commission intends to respond to the TCE report by consulting on a package of policy actions combining regulatory as well as market-led, co-regulatory options by 2015.

At the same time, future R&I actions under H2020 will address the optimal use and configuration of cloud computing solutions for data analytics and advanced infrastructures and services²⁴.

²² <http://ec.europa.eu/isa/>; http://ec.europa.eu/isa/documents/isa_lexuriserv_en.pdf

²³ 'Establishing a Trusted Cloud Europe: A policy vision document by the Steering Board of the European Cloud Partnership', <http://ec.europa.eu/digital-agenda/en/news/trusted-cloud-europe>

²⁴ See already, e.g., <http://www.helix-nebula.eu/>, a Cloud for Science initiative of ESA, CERN, EMBL etc.

2. E-infrastructures and High Performance Computing

Access to the best supercomputing facilities and services for industry, SMEs and academia is already provided by PRACE²⁵, a world-class HPC infrastructure for research.

Future actions will include establishing centres of excellence in the application of HPC to tackle scientific, industrial or societal challenges through the existing cPPP on HPC²⁶. Support will also be given to the development of the next generation of HPC technologies as a key horizontal enabler for advanced modelling, simulation and big data applications.²⁷

Given the urgent need to provide thousands of researchers across Europe with computing capabilities, key enablers of the digital European Research Area (ERA)²⁸, such as the European Grid Initiative, will also be supported.

3. Networks/ Broadband /5G

The on-going cPPP on 5G²⁹ works on the technological underpinnings of the mobile internet of the future and accompanies regulatory and funding initiatives³⁰ designed to encourage private investment in broadband infrastructure; It supports the evolution of backbone networks by increasing their capacity to handle large amounts of data.

4. Internet of Things (IoT)³¹

A series of large-scale projects will be funded to tackle the emerging questions of availability, quality and interoperability related to data gathered through smart connected objects and other IoT technology.

5. Public Data Infrastructures

The Commission will seek the Member States' support for an interconnected network of data processing facilities, through the linking of regional data centres and enabling infrastructures, to benefit from synergies and enhance efficiency, in particular for SMEs, academic, research organisations and the public sector. By reinforcing the GÉANT network³², the Commission will also invest in linking to non-EU countries, especially in the developing world.

²⁵ <http://www.prace-ri.eu/>

²⁶ http://ec.europa.eu/research/press/2013/pdf/ppp/hpc_factsheet.pdf

²⁷ <http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/calls/h2020-fethpc-2014.html>

²⁸ A Reinforced European Research Area Partnership for Excellence and Growth, COM(2012) 392.

²⁹ <http://5g-ppp.eu/>

³⁰ Such as the Connected Continent package (<http://ec.europa.eu/digital-agenda/en/connected-continent-single-telecom-market-growth-jobs>) and the telecommunications part of the Connecting Europe Facility (<http://ec.europa.eu/digital-agenda/en/connecting-europe-facility>).

³¹ The IoT is a dynamic global network infrastructure where physical and virtual "things" of all types communicate and are seamlessly integrated.

³² Pan-European data network for the research and education community, <http://www.geant.net/>.

4.2.3. Regulatory issues

1. Personal data protection and consumer protection

The fundamental right to personal data protection applies to big data where it is personal: data processing has to comply with all applicable data protection rules.

The Commission's reform package aims to build a single, modern, strong, consistent and comprehensive data protection framework for the EU. By strengthening individuals' trust and confidence in the digital environment and enhancing legal certainty, it will provide a regulatory environment essential for the development of innovative and sustainable data goods and services.

After adoption of the reform package, the Commission will work with Member States and stakeholders to ensure that business, and in particular SMEs, receive adequate guidance, notably on issues such as data anonymisation and pseudonymisation, data minimisation, personal data risk analysis, and tools and initiatives enhancing consumer awareness. The Commission will also actively support R&I for related technical solutions that are privacy-enhancing 'by design'.

On this basis, digital tools have the potential to help users to better control and secure their data. The Commission will launch a consultation process on the concept of user-controlled cloud-based technologies for storage and use of personal data ("personal data spaces"), and support R&I on tools to assist users in selecting the data sharing policies that best match their needs. It will also support projects aiming at reducing personal data breaches and ensuring that data is used in ways compatible with those for which it was collected.

Horizontal consumer and marketing law also applies to products based on big data technology. The Commission will ensure that SMEs and consumers, suppliers and users, are given all necessary information, are not misled, can rely on fair contracts, notably as regards the use of data collected from them. These measures will build the trust that is necessary to exploit the full potential of the data-driven economy.

2. Data-mining

The Commission is investigating ways in which data-driven innovation based on data-mining, including text-mining, might be enhanced, including in relation to the relevant copyright aspects.

The Commission takes note of Member States' initiatives that facilitate these activities by implementing (or reviewing the implementation of) the exceptions available under the current copyright framework.

3. Security

The Commission will explore the landscape of security risks relating to big data and will propose risk management and mitigation measures, including guidelines, e.g. on good practices for secure data storage, to further a security culture in many sectors of society and help detect and better respond to cyber-attacks.

The Commission will also support R&I to help reduce the risk of data breaches and of databases being exploited covertly for unlawful purposes.

4. Ownership/transfer of data

In several sectors, data location requirements limit the cross-border flow of information and form a barrier to a single market for cloud computing and big data. The Commission will study such barriers and will consider future policy actions, notably by taking into account the Trusted Cloud Europe report and recommendations put forward by the European Cloud Partnership.

In addition, the Commission will launch a consultation and expert group to assess the need for guidance on specific issues of data ownership and liability of data provision, in particular for data gathered through IoT technology.

5. Conclusions

A thriving data-driven economy will contribute to the well-being of citizens as well as to socio-economic progress through new business opportunities and through more innovative public services. It will flourish in a European Digital Single Market governed by modern and innovative rules.

The envisaged actions, once implemented, will result in accelerated innovation, productivity growth and increased competitiveness in data across the whole economy as well as on the global market with Europe as a key player.

The Commission will further consult with Parliament, Council, Member States and all relevant stakeholders to draw up a more detailed, multi-layered and evidence-based action plan for advancing towards the data-driven economy of the future and addressing Europe's future societal challenges.