



Brussels, 19.4.2016
COM(2016) 178 final

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

**European Cloud Initiative - Building a competitive data and knowledge economy in
Europe**

{SWD(2016) 106 final}
{SWD(2016) 107 final}

Introduction

The world is witnessing a dramatic increase in the amount and variety of data being produced. Alongside the data created by billions of people using digital devices and services for personal and professional reasons, and the data generated by the increasing number of connected objects, there is data from research, from digitised literature & archives and from public services such as hospitals and land registries. This "Big Data" phenomenon creates new possibilities to share knowledge, to carry out research and to develop and implement public policies.

It is also becoming easier to exploit this data thanks to the Cloud. The Cloud can be understood as the combination of three interdependent elements: the data infrastructures which store and manage data; the high-bandwidth networks which transport data; and the ever more powerful computers which can be used to process the data. The ability to analyse and exploit this Big Data is having an impact on the global economy and society, opening up the possibility of major industrial and social innovations. A key part of this impact is the change in the way scientific research is carried out, as we move rapidly towards [Open Science](#).

The Cloud makes it possible to move, share and re-use data seamlessly across global markets and borders, and among institutions and research disciplines. With the current capacity available in Europe, the data produced by EU research and industry is often processed elsewhere and European researchers and innovators tend to move to the places where high data and computing capacity is more immediately available. At the same time, as Europe is the largest producer of scientific knowledge in the world, it is well placed to take the global lead in the developing of a science cloud.

To fully exploit the potential of data as a key driver of Open Science and the [4th industrial revolution](#), Europe needs to answer several specific questions:

- How to maximise the incentives for sharing data and to increase the capacity to exploit them?
- How to ensure that data can be used as widely as possible, across scientific disciplines and between the public and the private sector?
- How better to interconnect the existing and the new data infrastructures across Europe?
- How best to coordinate the support available to European data infrastructures as they move towards exascale computing¹?

The potential gains for science, technology, and innovation from addressing these challenges were highlighted by the [scientific community itself](#), but also by [OECD Governments](#). The importance for the whole economy and society of addressing these challenges was confirmed by the EU's Member States in 2015.² This Communication proposes as a direct response a European Cloud Initiative which can secure Europe's place in the global data-driven economy.³

¹ Exa-scale computing refers to computing systems capable of at least one exaFLOPS – 10^{18} calculations per second – about 1.000 times faster than today's machines.

² See the [Conclusions](#) of the Competitiveness Council, 2015

³ Speech of President Juncker on October 2015; <http://bit.ly/1Y52pGi>

The European Cloud Initiative builds on the Digital Single Market (DSM) Strategy, which aims, *inter alia*, to maximise the growth potential of the European digital economy.⁴ It aims to develop a trusted, open environment for the scientific community for storing, sharing and re-using scientific data and results, the **European Open Science Cloud**.⁵ It aims to deploy the underpinning super-computing capacity, the fast connectivity and the high-capacity cloud solutions they need via a **European Data Infrastructure**.⁶ Focussing initially on the scientific community, the user base will be expanded to the public sector and to industry, creating solutions and technologies that will benefit all areas of the economy and society. Achieving this will require a collaborative effort open to all those interested in exploiting the data revolution in Europe as an essential component of global growth.

The European Cloud Initiative builds on the achievements of the European Cloud Strategy⁷ and the High Performance Computing (HPC) Strategy.⁸ It will build on initiatives such as the recently announced Important Project of Common European Interest (IPCEI) on HPC and Big Data enabled applications.⁹ It takes forward the policy developed in the Communication on Big Data¹⁰ and supports the European Open Science policy agenda which aims to increase the quality and impact of science,¹¹ building on the achievements of Open Access.¹² This Communication marks the beginning of a process whereby the Commission will engage with the Member States and with all relevant stakeholders to ensure the European Cloud Initiative can achieve its objectives.

The European Cloud Initiative will be complemented by further action under the Digital Single Market strategy covering cloud contracts for business users and switching of cloud services providers, as well as by the Free Flow of Data initiative.¹³

Five reasons why Europe is not yet fully tapping into the potential of data

First, many European businesses, research communities and public bodies are yet to tap into the full potential of **data** and of its potentially transformative effect on traditional sectors and on the way research is conducted.¹⁴ **Data coming from publicly funded research is not always open**; likewise data generated or collected by businesses is often not shared, and not always for commercial reasons. While some still see data as an asset to be protected many in business (especially SMEs), academia and public sectors are simply unaware of the value of data sharing. Reasons include **the lack of a clear structure of incentives** and rewards for

⁴ COM(2015) 192 final

⁵ Preparatory work started through a Commission High Level Expert Group, tasked to issue advice on its set-up: <http://bit.ly/1RK7lhh>

⁶ Preparatory work undertaken i.e. through advisory groups such as the e-infrastructures Reflection Group.

⁷ COM(2012) 529 final and results of working groups <http://bit.ly/1QVrvIb>

⁸ COM(2012) 45 final

⁹ The objective is to support the development of new industrial uses of HPC and to guarantee access to HPC facilities for public and private research, <http://bit.ly/1RMFq0i>

¹⁰ COM(2014) 442 final

¹¹ Council policy debate (9385/15); Council Conclusions (8970/15)

¹² COM(2012) 401 final

¹³ Possible legislative proposals will be subject to Commission better regulation requirements, in line with Commission's Better Regulation Guidelines, SWD(2015) 111

¹⁴ This is the case in e.g. health <http://bit.ly/1XEeaTN> (and ERC projects BIOTENSORS, DIOCLES, SMAC), astronomy (e.g. SparseAstro); climate change, migration or the Internet (e.g. DIADEM, MIGRANT, RAPID, THINKBIG).

data sharing (mainly in academia), of a clear legal basis¹⁵ (mainly in the public sector) and the shortage of data-related skills and lack of recognition of their value (in all sectors). The EU data protection framework prevents restrictions to the free movement of personal data on the grounds of privacy and personal data protection. Other legal and technical barriers to the free movement of data are yet to be tackled by the upcoming DSM initiative on the Free Flow of Data.

Second, **lack of interoperability** prevents addressing grand societal challenges that require efficient data sharing and a multidisciplinary, multi-actor approach, e.g. climate change which cannot be addressed by climatologists alone. While interoperability and data sharing have been tackled in some sectors (e.g. location of data by the INSPIRE Directive,¹⁶ health data by the patients' rights Directive¹⁷) many data sets remain unavailable to scientists, industry, public administrations and policy makers. While interoperability of administrative data mainly requires minimum standards, legal certainty in terms of access and usage and practical support,¹⁸ sharing research data is also hampered by the size of datasets, its varied formats, the complexity of the software needed to analyse it and deep-rooted walls between disciplines. Simple 'meta-data'¹⁹ to identify data and specifications for data-sharing are needed to make them widely accessible and available to be processed through common, open source data analysis tools. Issues of long-term preservation and curation of data must also be tackled. There are already global grassroots initiatives²⁰ and some Member States are advancing in this area but the European participation in these initiatives is limited and those efforts are largely fragmented.

Third, **fragmentation** hampers data-driven science.²¹ Data infrastructures are split by scientific and economic domains, by countries and by governance models. Access policies for networking, data storage and computing differ. Disconnected and slow data and computing infrastructures hinder scientific discovery, create silos and slow down the circulation of knowledge. Shareable research data, open data analysis tools and connected computing facilities need to be available to the great majority of researchers²² in Europe, not only to top scientists of leading disciplines from key research institutions. Moreover, Europe's universities and research centres generally operate within national structures and lack a European-scale environment for computational, storage and data analysis. This makes scientific cooperation in the EU more difficult, particularly multi-disciplinary cooperation based on data.²³ In a recent public consultation,²⁴ the vast majority of the respondents

¹⁵ The INSPIRE Directive 2007/2/EC provides an acquis for the sharing of European location data. However, the scope of the application of these laws is limited to specific data and services for environmental, natural disasters and health policies and not all obstacles with regard to data policies have been lifted effectively.

¹⁶ Regulation 1089/2010 implementing Directive 2007/2/EC

¹⁷ Work on the eHealth Network set up under Directive 2011/24 on patients' rights, eHealth Digital Service Infrastructure ePrescription and Patient Summary services to exchange of health data and the recent Joint Action supporting the eHealth Network report on the "Use of cloud computing in health" to support use of data other than for the direct care of an individual patient.

¹⁸ Tackled by the Commission's ISA Programme: <http://bit.ly/24DxWUs>

¹⁹ This may include high-quality statistical meta-data from official statistics to enhance data browsability, interoperability and integration.

²⁰ Several global initiatives address this: the FAIR data principles, the G8 Principles for Open Research Data Science, RDA guidelines, Belmont Forum recommendations, OECD principles and discipline-specific guidelines.

²¹ Consultation on Science 2.0 flagged the lack of integration of existing infrastructures as an obstacle for scientists' work.

²² Researchers are either unaware (54 %) or do not have a facility (37 %) to store and maintain their data (bit.ly/206u6hm).

²³ <http://bit.ly/1SkL9wm>

answered that the European Open Science Cloud would make science more efficient by better sharing of resources at national and international level.

Fourth, there is surging demand in Europe for **a world-class High Performance Computing (HPC) infrastructure to process data**²⁵ in science and engineering. The simulation of a complete next-generation airplane; climate modelling; linking genome to health; understanding the human brain; *in silico* testing of cosmetics to reduce animal testing – all need exascale computing capabilities. While in the long term, **quantum computing** holds the promise to solve computational problems that are beyond current supercomputers,²⁶ EU competitiveness also depends on the support of HPC for pan-European data infrastructures.

Worldwide, the USA, China, Japan, Russia and India are advancing swiftly. They have declared HPC a strategic priority, they fund programmes to develop national HPC ecosystems (hardware, software, applications, skills, services and interconnections) and work on the deployment of exascale supercomputers.²⁷ Europe is not participating in the HPC race in line with its economic and knowledge potential; it is falling behind other regions as it fails to invest in its HPC ecosystem and to reap the benefits of intellectual property in this field. On the supply side, EU industry provides about 5% of HPC resources worldwide, while it consumes one third of it. As Europe depends ever more on other regions for critical technology, the risk is that it gets technologically locked, delayed or deprived of strategic know-how. Europe also lags in terms of sheer total computing power: only one out of ten leading HPC infrastructures is in the EU, Germany's Höchstleistungsrechenzentrum Stuttgart ranking 8th. The USA has five and China has the fastest world supercomputer since 2013.

No single Member State alone has the financial resources to develop **the necessary HPC ecosystem**, in competitive time frames with respect to the US, Japan or China.²⁸ However to date no common action is taken to bridge the gap between internal demand and EU supply.²⁹ The EU set up a contractual Public-Private Partnership on HPC to develop exascale technology, but there is no European framework to integrate it in large scale computing systems.

Finally, scientific data producers and users must be able to re-use data and to use advanced analytics techniques, such as text and data mining, in an environment that is at least as dependable as their own facilities. Member States have made strong reference to the importance of EU research data and to ensure that data-driven science benefits European society.³⁰ Any use and re-use of scientific data needs to ensure that personal data are adequately protected according to the EU data protection rules.³¹ These and forthcoming revision of EU Copyright legislation³² provide general frameworks which are relevant in this context.

²⁴ <http://bit.ly/1JEymCY>

²⁵ Requests for computing cycles about double PRACE availability: <http://bit.ly/1So2sgc>

²⁶ SWD(2016) 107

²⁷ SWD(2016) 106

²⁸ The US Department of Defence will invest \$525 Million for acquiring 3 pre-exa-scale systems in 2017/2018 ('CORAL'). Japan plans to invest \$1.38 Billion to install a near-exa-scale system in 2019.

²⁹ While PRACE allows sharing computing resources of some Member States, the procurement of HPC systems is a national decision without EU coordination or funding.

³⁰ Council conclusions (8970/15).

³¹ COM (2012) 9 final

³² COM (2015) 626 final

What are the solutions?

1. European Open Science Cloud

The European Open Science Cloud aims to give Europe a global lead in scientific data infrastructures, to ensure that European scientists reap the full benefits of data-driven science. Practically, it will offer 1.7 million European researchers and 70 million professionals in science and technology a virtual environment with free at the point of use, open and seamless services for storage, management, analysis and re-use of research data, across borders and scientific disciplines. Its development will be driven by the scientific community, who are the most advanced users and the largest producers of science in the world. The European Open Science Cloud will be also open for education and training purposes in higher education and, over time, to government and business users as the technologies developed will be promoted for wider application.

The European Open Science Cloud will start by federating existing scientific data infrastructures, today scattered across disciplines and Member States. This will make access to scientific data easier, cheaper and more efficient. It will enable the creation of new market opportunities and new solutions in key areas such as health, environment, or transport. The European Open Science Cloud will provide a secure environment where privacy and data protection must be guaranteed by design, based on recognised standards, and where users can be confident concerning data security and liability risks. It will leverage other actions taken by the Commission to promote Open Science in Europe, such as open access to scientific publications and data in Horizon 2020, and convene key stakeholders to co-design the next actions. The governance of the European Open Science Cloud will be determined upon conclusion of a thorough preparation process that is already under way.

Specifically, to develop the European Open Science Cloud it will be necessary to:

- **Make all scientific data produced by the Horizon 2020 Programme open by default.** This will extend the current pilot,³³ whereby projects implement Data Management Plans to make research data findable, accessible, interoperable and re-usable (FAIR principles).³⁴
- **Raise awareness and change incentive structures** for academics, industry and public services to share their data, and improve data management training, literacy and data stewardship skills. In parallel, principles and guidelines on access to research data in Europe³⁵ will be reviewed to strengthen and coordinate their implementation.

³³ The Horizon 2020 Open Research Data Pilot currently covers : Future and Emerging Technologies, Research infrastructures, Information and Communication Technologies, the ‘nanosafety’ and ‘modelling’ topics in the Nanotechnologies, Advanced Materials, Advanced Manufacturing and Processing, and Biotechnology WP, selected topics within the Societal Challenges: Food security, sustainable agriculture and forestry, marine and maritime and inland water research and the bioeconomy; Climate Action, Environment, Resource Efficiency and Raw materials – except raw materials; Europe in a changing world – inclusive, innovative and reflective Societies; Science with and for Society as well as the cross cutting activity and focus area Smart and Sustainable Cities. Note that projects that are not part of these "core areas" can still join on a voluntary basis.

³⁴ The existing opt-out options, where open access to data would be contrary to future commercial application or data privacy and personal data protection, security and protection of EU classified information will be maintained. The analysis of the pilot showed that most projects apply open data, but that opt-out options are also important.

³⁵ C(2012) 4890 final

- Develop specifications for **interoperability and data sharing** across disciplines and infrastructures, building on existing initiatives such as the Research Data Alliance and the Belmont Forum and legal provisions such as INSPIRE. Over time, any emerging standardization needs will be addressed through the DSM Priorities for ICT Standardisation.
- Create a **fit-for-purpose pan-European governance structure** to federate scientific data infrastructures and overcome fragmentation. The institutional set-up will oversee long-term funding, sustainability, data preservation and stewardship. It will build on existing structures to involve scientific users, research funders and implementers.³⁶
- Develop **cloud-based services for Open Science**. Supported by the European Data Infrastructure, they will allow researchers to find and access shared research data, to employ advanced analytical software, to use high-performance computing resources and to learn about best data-driven science practices from leading disciplines.
- **Enlarge the scientific user base** of the European Open Science Cloud to researchers and innovators from all disciplines and Member States, as well as from partner countries and global initiatives, so that they contribute to excellence and partake in the benefits of the initiative.³⁷

The initiative will reinforce other Open Science actions that the Council³⁸ and the European Parliament³⁹ called for, and actions in the context of the forthcoming Open Science policy agenda of the Commission. It will foster best practices of data findability and accessibility and help researchers get their data skills recognised and rewarded; allow easier replicability of results and limit data wastage e.g. of clinical trial data (research integrity); contribute to clarification of the funding model for data generation and preservation, reducing rent-seeking and priming the market for innovative research services (e.g. advanced text and data mining). The initiative may also help address issues of data clearance and personal data protection.⁴⁰ The Commission will consult stakeholders and work with R&D providers on the need for implementing guidelines for the scientific domain regarding Union policy and law in Data Protection, and on the need to ensure that the initiative implements "by-design" the legal principles at the earliest possible stage.

Actions	Timeline
The Commission will work with global policy and research partners to foster cooperation and to create a level playing field in scientific data sharing and data-driven science.	As of 2016

³⁶ Such as ESFRI, INSPIRE, eIRG, GEANT, PRACE, ELIXIR, the Belmont Forum and similar federating initiatives.

³⁷ New Commission initiatives may be financed under ESIF provided that the Member States agree to finance them and to modify their operational programmes accordingly.

³⁸ Council Conclusions (8970/15).

³⁹ European Parliament Report 2015/2147(INI).

⁴⁰ While fully respecting Articles 7 and 8 of the Charter of Fundamental Rights of the European Union and current and forthcoming provisions on use of data for research purposes, the initiative may develop, for instance, services for IPR-conscious text and data mining, access control for different uses, irreversible anonymisation of sensitive data before data fusion, 'personal data spaces' to preserve privacy and foster uptake of innovative uses or rely on machine-readable licencing and privacy meta-data attached to data-sets accessible via the cloud and provide guidelines and best practices for compliant organisational processes supporting the initiative. While these are technical, by-design and by default tools and processes, they may help reduce the incidence of malpractice and reduce non-conformity with legal provisions.

The Commission will use the Horizon 2020 Work Programmes to provide funding to integrate and consolidate e-infrastructure platforms, to federate existing research infrastructures and scientific clouds and to support the development of cloud-based services for Open Science.	As of 2016
The Commission will make open research data the default option, while ensuring opt-outs, for all new projects of the Horizon 2020 programme.	As of 2017
The Commission will review the 2012 Commission Recommendation on access to and preservation of scientific information ⁴¹ to encourage scientific data sharing and the creation of incentive schemes, rewards systems and education and training programmes for researchers and businesses to share data, in close relation with the DSM 'Free flow of data' initiative.	As of 2017
The Commission will work with Member States to connect the priority European research infrastructures ⁴² to the European Open Science Cloud.	As of 2017
Together with stakeholders and relevant global initiatives, the Commission will work towards an Action Plan for scientific data interoperability, including 'meta-data', specifications and certification.	By end 2017

2. European Data Infrastructure

The **European Data Infrastructure**, once fully implemented, will underpin the European Open Science Cloud. Europe needs integrated world-class HPC capability, high-speed connectivity and leading-edge data and software services⁴³ for its scientists and for other lead users from industry (including SMEs) and the public sector. This infrastructure will allow fully unlocking the value of Big Data and digital by default.⁴⁴ The European Data Infrastructure will also support the EU to rank among the world's top supercomputing powers by **realising exascale supercomputers around 2022, based on EU technology, which would rank in the first 3 places of the world**. Europe should aim to have at least two sources of this technology.

While the existing HPC strategy⁴⁵ supports research and development of marketable HPC technologies, it does not foresee realising an exascale supercomputer. The European Data Infrastructure will gather the necessary resources and capabilities, to close the chain from research and development to the delivery and operation of the exascale HPC systems co-designed between users and suppliers. This will include data connectivity and big data storage to make sure that supercomputing services are available across the EU, no matter where supercomputers are located. A first step was recently taken by Luxembourg, France, Italy and Spain, with an **Important Project of Common European Interest (IPCEI) on HPC and Big Data enabled applications**.⁴⁶

Building on the Pan-European High Performance Computing infrastructure and services

⁴¹ C(2012) 4890 final

⁴² As identified by the European Strategic Forum on Research Infrastructures (ESFRI) <http://bit.ly/1pfqOe7>

⁴³ Including existing services by OpenAIRE, EUDAT, EGI, IndigoDataCloud, HelixNebula, PRACE, GÉANT.

⁴⁴ 'Digital by default' refers to services and processes that are made available online or in a digital form by default.

⁴⁵ COM(2012) 45 final

⁴⁷ <http://bit.ly/1QxERan>

(PRACE), the trans-European high speed network (GÉANT), the contractual Public-Private Partnership on HPC,⁴⁷ the ECSEL Joint Undertaking,⁴⁸ and the IPCEI on HPC and Big Data, the Commission and participating Member States will:

- foster an HPC ecosystem capable of developing new European technology such as **low power HPC chips**;⁴⁹
- integrate technologies into system prototypes, co-designing⁵⁰ solutions and procuring HPC systems; the resulting HPC infrastructure will focus on **supercomputers of top-range capabilities** connected to mid-range EU national computing centres and to pan-European data and software infrastructure to offer supercomputing as a service;
- provide **seamless, high-speed, reliable and secure connectivity** to make HPC accessible across the EU; the trans-European high speed network (GÉANT) and National Research and Education Networks (NREN) already connect 50 million researchers and students; these infrastructures will be upgraded to match the increase of data volumes to be transferred and the extension of the user base.

The **European Data Infrastructure** will contribute to the digitisation of industry, to develop European platforms for new, strategic applications (e.g. medical research, aerospace, energy) and to foster industrial innovation. It will **widen the user base of HPC**, providing easier access via the Cloud both to researchers in key scientific disciplines and to the long tail of science. Industry, particularly SMEs without in-house capabilities and public authorities (e.g. smart cities and transport) will benefit from cloud-based and easy-to-use HPC resources, applications and analytics tools.⁵¹ In this context, the Commission will foster the deployment of processing and exploitation capacities for Sentinel satellites' data, Copernicus services information and other Earth Observation data, so as to enable the cross-fertilisation of different data sets, encourage the development of innovative products and services and maximise the socio economic benefits of Earth Observation data in Europe.

The European Data Infrastructure will work in combination with the national and regional, scientific and public data centres. It will develop and implement best practices based on certification schemes, common European and global standards and specifications⁵² to tackle the current lack of interoperability between national and disciplinary data centres.⁵³

The European Data Infrastructure will include a governance structure for the management and the development of the data infrastructure and services,⁵⁴ decision making on funding, long-term sustainability and security. The governance should involve users (the European Open

⁴⁷ <http://bit.ly/1WZH8wF>

⁴⁸ <http://www.ecsel-ju.eu>

⁴⁹ Energy efficient exa-scale machines would impact the full spectrum of computing and bring Europe technical, economic and social advantages. Currently, a single exa-scale machine would need a dedicated power plant of 700MW to operate, enough to power 140.000 households for a year. Thus the need for low power chips.

⁵⁰ Co-design is an approach to design aiming to actively involve customers and users in the design process to help ensure the result meets their needs and is usable.

⁵¹ <http://bit.ly/1pqny20>

⁵² RDA-Europe has started the interaction with the ICT standardisation multi stakeholder group to submit best practice implementations for data infrastructure interoperability developed in the Research Data Alliance.

⁵³ Such as INSPIRE interoperable spatial data and services specifications.

⁵⁴ Building on existing services by OpenAIRE, EUDAT, EGI, IndigoDataCloud, HelixNebula, PRACE, GÉANT.

Science Cloud and other long-term users such as the public sector), implementers (PRACE, GEANT) and funders, and should build on existing governance structures.

Actions	Timeline
<p>The Commission and participating Member States should develop and deploy a large scale European HPC, data and network infrastructure, including:</p> <ul style="list-style-type: none"> – the acquisition of two co-designed, prototype exascale supercomputers and two operational systems which will rank in the top three of the world; – the establishment of a European Big Data centre,⁵⁵ – the upgrade of the backbone network for research and innovation (GEANT) and the integration of European public services networks. 	<p>2016-2020</p> <p>as of 2018</p> <p>as of 2016</p> <p>as of 2016</p>

Exploiting the potential of quantum technologies

The next breakthrough in supercomputing and secure networking may be based on quantum technologies. Leading companies in Europe, Asia-Pacific and North America are starting to invest in quantum, but a higher level of investment is necessary to reach marketable products. Europe has to be at the forefront of these future advances.⁵⁶ The European Data Infrastructure should be complemented by an ambitious, long-term and large-scale flagship initiative to unlock the full potential of quantum technologies, accelerate their development and bring commercial products to public and private users. The European Commission will start the preparatory steps for the flagship, including consultation of stakeholders, impact assessment, taking into account results of the interim evaluation of the Horizon 2020 Programme by the end of 2017.⁵⁷

Action	Timeline
<p>The European Commission will start the preparatory steps for the flagship, including consultation of stakeholders, impact assessment,⁵⁸ taking into account results of the interim evaluation of the Horizon 2020 Programme by the end of 2017,⁵⁹ with the aim to launch the ramp up phase in 2018.⁶⁰</p>	<p>2016-2019</p>

3. Widening access and building trust

The uptake of cloud services in the public sector is uneven and slow.⁶¹ This is due to lack of trust and limited synergies between the public sector and the academia. Fragmentation in terms of data infrastructures is an obstacle for building critical mass and common solutions for different user groups. **The user base of the European Open Science Cloud and of the**

⁵⁵ E.g. hosted by JRC for multidisciplinary data but focused on INSPIRE/GEOSS/Copernicus spatial data.

⁵⁶ <https://goo.gl/zBVi8N>

⁵⁷ SWD(2016) 107

⁵⁸ The impact assessment will be part of the preparatory process for relevant funding programmes under the post 2020 financial perspective. Any additional implementation measure likely to have significant impacts may require separate, individual impact assessments.

⁵⁹ SWD(2016) 107

⁶⁰ FET Flagships as described in Horizon 2020 reference documents.

⁶¹ SMART 2013/0043: Public sector organisations trail the private sector, with a 10% difference in 2013 in the use of cloud computing services.

European Data Infrastructure will be widened to the public sector, for example through large-scale pilots involving eGovernment⁶² and public sector stakeholders and by progressively opening the European Data Infrastructure to **users from industry** and the public sector to achieve a European dimension. Over time, the European Open Science Cloud will ensure that public data is fully discoverable, accessible and exploitable by scientists, policy makers and businesses. Lessons learnt will provide concrete guidance for the adoption of cloud-based services by public administrations across Europe.

As the public sector generates massive amounts of data (e.g. Copernicus earth observation, INSPIRE location data) and needs larger computing capacity (e.g. for real time traffic and travel information systems, for smart city applications or for policy modelling), it will benefit from economies of scale, flexibility and continuity. The public will thus benefit from cheaper, faster, better and interconnected public services and from better policy making based on affordable and secure compute- and data-intensive services.

In similar fashion, the European Open Science Cloud and the European Data Infrastructure will benefit businesses, including SMEs, that lack cost-effective and easy access to data storage, services and advanced computing. Actions will be taken to progressively widen the user base to innovative SMEs and industry, via data and software centres of excellence and data services innovation hubs for SMEs. These actions will require close cooperation with the private sector: SMEs, large scientific and industrial users of HPC and the Cloud services industry, who needs to be involved from the start.

Additionally, the European Cloud Initiative will need to meet **high standards of quality, reliability and confidentiality**, to ensure protection of personal data and intellectual property, **and security** – in terms of resilience and protection against intrusion. Existing public sector facilities - notably the Connecting Europe Facility (CEF) Digital Service Infrastructure (DSI) building blocks related to trust and security - can be reused and deployed by the scientific community for cost savings, ease of access, and overall consistency. The general framework will be provided by the general data protection rules, the NIS Directive⁶³ and the revision of the EU copyright legislation. Given the global nature of cloud computing, it is essential that the European data economy remains connected to the rest of the world and that the global standards of data protection are raised to a high level essentially equivalent to that in Europe.

Working on appropriate standards is part of the DSM Priorities for ICT Standardisation Plan⁶⁴; a suitable certification scheme will be designed at EU level to guarantee security, data portability, and interoperability in compliance with legal requirements,⁶⁵ including the certification scheme already provided for in General Data Protection Regulation for personal data security. While a number of certification schemes⁶⁶ exist, their scope and application vary considerably, and there is no common approach on minimum requirements in the procurement or management of public sector cloud resources. In this respect, collaboration with industry and public authorities will match the capability of industry with the requirements of science and the public sector.

⁶² EU eGovernment Action Plan 2016-2020 – accelerating the digital transformation of government

⁶³ COM (2013) 48

⁶⁴ COM(2016) 176

⁶⁵ Regulation 765/2008

⁶⁶ <https://resilience.enisa.europa.eu/cloud-computing-certification>

Widening of access to the European Open Science Cloud and European Data Infrastructure will be carried out in line with the appropriate legislation, in particular for what concerns the re-use of data for other purposes.

Actions	Timeline
<p>In partnership with industry and the public sector, the Commission undertakes to:</p> <ul style="list-style-type: none"> - adapt HPC and Big Data solutions to a cloud environment in order to enable broad access, notably for SMEs; - develop an ecosystem to strengthen the cloud industry in Europe, using the European Open Science Cloud as a testbed for innovative cloud technology solutions; - create a platform for public authorities to open their data and services, creating a "Government as a Service" (GaaS) base for the EU. 	2016-20
<p>In order to facilitate the uptake of big data technologies, the Commission will provide a big data test environment (large-scale pilots) for public administrations, including in the framework of the proposed IPCEI.</p>	As of 2016
<p>The Commission working with industry and Member States will promote the use of existing relevant certifications and standards, and – where appropriate – the creation of European-level certification and labelling, in particular to support public procurement of cloud services.</p>	As of 2016

Financial implications

Digital transformation in Europe requires scale. Various sources of EU financing can be identified for the European Cloud Initiative:

- Horizon 2020 Framework Programme for Research and Innovation (Horizon 2020)
- Connecting Europe Facility (CEF)
- European Structural and Investment Funds (ESIF)
- European Fund for Strategic Investments (EFSI)

Different sources of financing are necessary to support the full investment cycle. Large infrastructure projects are supported initially by public grants and, as they mature, by risk-sharing and market-based instruments. However, since such initiatives require consistent and coordinated efforts, the fragmentation of available budget sources is clearly a drawback.

Existing funding under Horizon 2020 will allow to support the European Open Science Cloud and to kick-start the European Data Infrastructure. Initial estimation of the required additional public and private investment is €4.7 billion in the period of 5 years. This includes €3.5 billion for data infrastructure,⁶⁷ €1 billion for a large-scale EU-wide Quantum Technologies flagship and €0.2 billion for actions on widening access and building trust. Additional provisions will be discussed with Member States for enlarging support to the European Open Science Cloud beyond Horizon 2020. The initiative will over time generate revenue of its own as its use by the scientific community, innovative start-ups and the public sector takes off.

⁶⁷ SWD(2016) 106

The Commission intends to propose how the different sources of financing at EU and national level could be blended in order to realise the objectives of the present Communication in full; it will discuss them with the Member States following appropriate evaluation, assessments of impacts and consultation. Infrastructure of this level of ambition will require strong involvement of the Member States, in particular by leveraging structural funds and EFSI⁶⁸ guarantees, but also significant investments from the private sector and appropriate coordinating mechanisms. In this respect, the proposed Important Project of Common European Interest (IPCEI) on HPC and Big Data shows the possibilities and the positive effects of Member States' engagement.

Actions	Timeline
In cooperation with Member States and stakeholders, the Commission will explore appropriate governance and financing mechanisms for the Open Science Cloud and the European Data Infrastructure and define an implementation Roadmap.	As of 2016
The Commission will put forward approaches for blending different funding streams, for discussion with Member States and stakeholders, in order to realise the objectives of this Communication.	2016

CONCLUSIONS

The European Cloud Initiative is designed to help science, industry and public authorities in Europe access world-class data infrastructures and cloud-based services as they become the decisive factors for success in the digital economy.

A European Cloud Initiative should open up to every research centre, every research project and every researcher in Europe the world-class supercomputing, data storage and analysis capacity which they need to succeed in the global, data-driven innovation system.

The Initiative will make it possible to widen the user-base of the infrastructures and services to the public sector and to industry, including SMEs, guaranteeing an adequate level of security, data portability, interoperability as well as compliance with EU legal requirements.

The extent to which the Member States and private sector embrace the gains to be made from addressing this challenge, and commit to working together to address them will determine the success of the initiative.

⁶⁸ Advisory services of EIB under the European Investment Advisory Hub will also be involved.